

Möglichkeiten und Grenzen der spektralanalytischen Untersuchung ökonomischer Zeitreihen

DISSERTATION

vorgelegt dem Fachbereich
Rechts- und Wirtschaftswissenschaften
der Johannes Gutenberg-Universität Mainz
zur Erlangung des Grades eines
Doktors der Staatswissenschaften
(Dr. rer. pol.)

von
Diplom-Mathematiker
Eugen Oswald

1976



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41547917_0043

Möglichkeiten und Grenzen der spektralanalytischen Untersuchung ökonomischer Zeitreihen

DISSERTATION

vorgelegt dem Fachbereich
Rechts- und Wirtschaftswissenschaften
der Johannes Gutenberg-Universität Mainz
zur Erlangung des Grades eines
Doktors der Staatswissenschaften
(Dr. rer. pol.)

von
Diplom-Mathematiker
Eugen Oswald

Möglichkeiten und Grenzen
der spektralanalytischen Untersuchung
ökonomischer Zeitreihen

D I S S E R T A T I O N

zur Erlangung des Grades eines
Doktors der Naturwissenschaften
an der Universität Leiden
vorgelegt von
Dr. phil. habil. E. Härtter



Erstberichterstatter: Prof. Dr. H. Stöwe

Zweitberichterstatter: Prof. Dr. E. Härtter

Tag der mündlichen Prüfung: 20. Dezember 1976

Inhaltsverzeichnis

Einleitung	5
I. Methodische Gesichtspunkte zur Analyse ökonomischer Reihen	
1. Ökonomische Reihen und ihre Komponenten	9
1.1 Allgemeine Beschreibung	9
1.2 Ökonometrische Zeitreihenmodelle	11
1.3 Die Komponenten im Rahmen ökonometrischer Modelle	15
2. Zeitreihenanalyse als Bestimmung der Komponenten	19
3. Zwei methodische Beispiele	21
3.1 Ein Regressionsmodell	21
3.2 Die Methode der gleitenden Durchschnitte	23
4. Die Beurteilung eines Analyseverfahrens	25
4.1 Kriterien für das Analysemodell	25
4.2 Forderungen an die Schätzmethode	27
II. Wahrscheinlichkeitstheoretische Grundlagen stationärer Prozesse	
5. Stochastische Prozesse	31
6. Stationäre stochastische Prozesse	35
7. Spektraldarstellungen bei stationären Prozessen	39
7.1 Prozesse mit orthogonalen Zuwächsen	39
7.2 Komplexe Form der Spektraldarstellungen	43
7.3 Reelle Form der Spektraldarstellungen	46
8. Lineare Funktionen stationärer Prozesse	48
9. Beispiele	50

9.1	Der Prozeß aus identischen Zufallsvariablen	50
9.2	Prozesse mit konstantem Erwartungswert	51
9.3	Der harmonische Prozeß	52
9.4	Der stationäre Gaußsche Prozeß	52
9.5	Der white-noise Prozeß	53
9.6	Der gleitende Durchschnitt	54
9.7	Der autoregressive Prozeß	55
9.8	Ein Beispiel von Doob	58
9.9	Ein Beispiel von Nerlove	59
III. Lineare Prognose bei stationären Prozessen		
10.	Das Prognoseproblem	61
11.	Reguläre und deterministische Prozesse	64
12.	Die Zerlegung von Wold	68
IV. Statistische Grundlagen der Spektralanalyse		
13.	Ergodische Eigenschaften stationärer Prozesse	71
14.	Das Periodogramm	74
14.1	Wieners Periodogrammbegriff	74
14.2	Der Periodogrammbegriff bei stochastischen Prozessen	76
15.	Konsistente Schätzung der spektralen Dichte	78
15.1	Das gewichtete Periodogramm	78
15.2	Einige Spektralfenster	81
15.3	Prewithening und Recoloring	84
V. Spektralanalyse ökonomischer Reihen		
16.	Motivationen für den Ansatz	85
17.	Die Stationarität der Komponenten	88
17.1	Allgemeines	88
17.2	Der Trend	89
17.3	Die Saisonkomponenten	90

17.4	Die zyklische Komponente	92
18.	Die Ausschaltung von Komponenten	93
18.1	Filtermethoden	93
18.2	Interpretationsprobleme	95
19.	Schätzprobleme beim Spektrum	96
19.1	Fragen zur Technik	96
19.2	Beurteilung einiger Schätzergebnisse	98
VI.	Zusammenfassende Beurteilung und Konsequenzen	
20.	Kritik des Analysemodells	103
21.	Kritik der Ziele	106
22.	Kritik der Schätzmethoden	109
23.	Alternativen zur Spektralanalyse	111
23.1	Ein Regressionsmodell mit stationärem "Rest"	111
23.2	ARMA und ARIMA-Modelle	113
	Literaturverzeichnis	115

Einleitung

Im Gegensatz zu Untersuchungen bei Querschnittsdaten kann bei den Daten einer Zeitreihe im allgemeinen nicht davon ausgegangen werden, daß sie unabhängig voneinander sind. Diese typische Eigenschaft bildet zweifellos den Mittelpunkt der statistischen Zeitreihenanalyse. Spätestens seit dem Erscheinen der Aufsätze von Slutsky (45) und Yule (54) ist eine besondere Form der Abhängigkeit, nämlich die der periodischen Wiederholungen, verstärkt in den Vordergrund getreten und in statistische und wahrscheinlichkeitstheoretische Überlegungen mit einbezogen worden.

Schon das von Schuster entwickelte Periodogramm diente zur Auffindung "versteckter Perioden" in Zeitreihen.

Zu seiner Berechnung und Beurteilung bediente er sich aber noch der klassischen mathematischen Disziplin der Fourieranalyse, die mit der Wahrscheinlichkeitstheorie und speziell mit den stochastischen Prozessen in keiner Verbindung steht. Erst Wiener (51) und Khintchine (23) haben die formalen Grundlagen geliefert, die eine Interpretation der Fourierzerlegung der Autokorrelationsfunktion bestimmter Zeitreihen auf wahrscheinlichkeitstheoretischer Basis erlauben.

Die Einbettung der Reihenkorelation in den Rahmen der stationären stochastischen Prozesse erwies sich in der Folgezeit vor allem für die stochastische Zeitreihentheorie als richtungsweisend und fruchtbar. So entwickelten Kolmogorov (26) und wiederum Wiener (52) eine Prognosetheorie für stationäre Zeitreihen, in der insbesondere die Frage nach dem besten Prognosewert für einen zukünftigen Zeitpunkt, wenn die Vergangenheitswerte bekannt sind, gelöst wird. In der nach Wold (48) benannten Zerlegung eines stationären Prozesses wird die Prognosetheorie dazu benutzt, zwei additive Komponenten, die sogenannten

"deterministischen" und "regulären" Anteile des Prozesses zu erzeugen, die sich in Bezug auf die Prognose sehr verschieden verhalten.

Nimmt man zu diesen wichtigen Ergebnissen noch die auf Wieners und Khintchines Grundlagen zurückgehende sogenannte Spektralzerlegung hinzu, dann läßt sich feststellen, daß die wahrscheinlichkeitstheoretische Struktur der stationären Prozesse weitgehend geklärt ist.

Auch die statistische Seite der Untersuchungen hat sich stark fortentwickelt. Im Mittelpunkt steht die Spektralanalyse, d.h. die Bestimmung bzw. Schätzung der charakteristischen Größen der oben genannten Spektralzerlegung aus einer realisierten Zeitreihe. Das von Schuster entwickelte Periodogramm spielt dabei in einer etwas modifizierten Art wieder eine wichtige Rolle.

Von den zahlreichen Veröffentlichungen zu diesem Thema seien hier einige der ersten wie Bartlett (4), Grenander und Rosenblatt (17), Blackman und Tukey (5) sowie Parzen (37) genannt. Ein vergleichsweise geringer Teil der Literatur zu statistischen Problemen bei stationären Zeitreihen befaßt sich auch mit Testfragen zur Erkennung von bestimmten stationärem Verhalten bei realisierten Reihen. Neben Whittle (50) haben dazu auch Durbin und Watson (8), (9) Arbeiten geliefert, in denen es gelang, beim Test auf Unabhängigkeit der Restwerte eines Regressionsmodells den Einfluß der erklärenden Variablen auf die Schätzung der realisierten Restwerte weitgehend auszuschalten. Die Alternative zur Unabhängigkeit der Restwerte ist dabei ein autoregressiver Prozeß erster Ordnung, also ein spezieller stationärer Prozeß.

Die statistische und stochastische Theorie der stationären Prozesse ist sehr stark durch ein Hauptanwendungsgebiet, nämlich die Radio-, Elektro- und Nachrichtentechnik beeinflusst worden. Das bezeugen die technischen Begriffe wie Filter oder white noise

(weißes Rauschen), die in diesem Teilgebiet der Physik ihre Erklärung finden. Wie bei vielen anderen statistischen Methoden ist dies aber nicht der einzige Anwendungsbereich geblieben. Bedenkt man die zahlreichen periodischen Erscheinungen, die bei ökonomischen Reihen beobachtbar sind, dann ist es verständlich, daß auch zur Analyse von solchen Zeitreihen die Theorie stationärer Prozesse und insbesondere die Spektraltheorie herangezogen wurde. In Bezug auf die Saisonkomponente haben dies beispielsweise Hannan(19) und Nerlove (35) getan. Bei Granger und Hatanaka (14) werden eine ganze Reihe weiterer Anwendungen beschrieben.

Die Frage der Anwendbarkeit, d.h. die Frage, ob ökonomische Zeitreihen oder einige ihrer Komponenten als stationäre Prozesse aufgefaßt werden können, wird in der diesbezüglichen Literatur nur am Rande erwähnt. Eine negative Einstellung findet man beispielsweise bei Grenander und Rosenblatt (17, S. 116) :

" However, we believe that the results obtained indicate that a statistical theory of stationary time series of great practical value can be constructed. The success of such nonparametric approach in certain fields does not guarantee its success in other fields e.g. in the social sciences and economics. This can be judged only after application of these techniques to relevant data arising in these fields, but one suspects that the stationary assumption may break down in such contexts." Die hier geäußerte Vermutung wird allerdings nicht weiter untersucht. Überwiegend werden in den Veröffentlichungen die Methoden und Verfahren allerdings kritiklos dargestellt, und die Anwendbarkeit nicht diskutiert.

Nach meiner Ansicht bieten jedoch die Wahrscheinlichkeitstheorie auf der einen Seite und ökonomische Gesichtspunkte zur Zeitreihenanalyse auf der anderen Seite genügend Möglichkeiten, um zu die-

sem Problem definitive Antworten zu geben. Darin besteht das Anliegen der vorliegenden Arbeit. Es wird nachfolgend die These aufgestellt, daß im Bereich der univariaten Zeitreihenanalyse ökonomischer Reihen weder die Trendkomponente noch die Saisonkomponente den Voraussetzungen einer Spektralanalyse genügen, und daß die Spektralanalyse auch für die übrigen Komponenten ein unbefriedigendes Instrument ist.

Die These bedeutet nicht, daß stationäre Prozesse in der Analyse ökonomischer Reihen nicht verwandt werden könnten. Es wird vielmehr gezeigt, daß diese Modelle für die zyklische Komponente sehr sinnvoll sind. Allerdings benötigt man andere Formen als die der Spektraldarstellung. Dazu werden am Schluß der Arbeit Möglichkeiten aufgezeigt.

I. Methodische Gesichtspunkte zur Analyse ökonomischer Reihen

1. Ökonomische Reihen und ihre Komponenten

1.1. Allgemeine Beschreibung

Für eine Gliederung ökonomischer Zeitreihen gibt es mehrere Möglichkeiten. Zu ihnen gehören die Einteilung nach Bestands- und Flußgrößen, die Unterscheidung nach Mikro- und Makrovariablen, nach Real- und Nominalgrößen, die Zusammenfassung nach Sektoren oder auch Gruppenbildungen wie Preise, Löhne und Gehälter, Produktion oder Beschäftigung.

Jede dieser Einteilungen kann verschiedene Probleme aufwerfen. Bei Bestandsgrößen etwa tritt die Frage nach dem Verlauf zwischen den Beobachtungszeitpunkten auf, die i.a. nicht eindeutig beantwortet werden kann. Flußgrößen müssen unter Umständen normiert werden, wegen der verschiedenen Zeitspannen, die zwischen den Beobachtungspunkten liegen. Ein praktisches Beispiel ist die arbeits-tägliche Bereinigung von Produktionsdaten. Bei Mikrovariablen tritt die Frage nach der Aggregation auf, bei Nominalgrößen die Frage der Preisbereinigung.

Neben dem Streben nach Systematisierung ökonomischer Reihen ist auch das Anliegen vorhanden, für alle Reihen gleichermaßen gültige Charakteristiken zu finden. Schon früh wurde deshalb auch die Erscheinungsform einer Reihe in die Betrachtung einbezogen. Historisch interessant ist z.B., daß Kondratieff (27, S. 575) "alle Elemente der ökonomischen Wirklichkeit nach dem Charakter ihrer Dynamik in zwei Gruppen" einteilt. Die eine Gruppe, zu der Produktion und

Verbrauchs-größen zählen, ist dadurch gekennzeichnet, "daß sie,..., in ihrer Dynamik neben Schwankungsvorgängen auch eine allgemeine Tendenz von einer bestimmten Richtung - in der Regel nach oben - bekunden". Die andere Gruppe, zu der insbesondere die Preise gehören, weist eine solche Tendenz nicht auf. Das Ergebnis von Kondratieffs Untersuchungen waren die nach ihm benannten langen Wellen, die er besonders in Preis- und Produktionsreihen zu entdecken glaubte, deren Nachweis aber von Weinstock (49) angezweifelt wird. Bekannt sind weiterhin Kuznets' lange Wellen, die sogenannten Bauzyklen, sowie die Juglars und Kitchins bei Schumpeter, die nach einer "Definition" des National Bureau of Economic Research (14, S. 16) zu den "business cycles" gezählt werden können.

Spätestens seit der Konjunkturforschung der zwanziger und dreißiger Jahre ist es üblich, eine Zeitreihe in solche zyklischen Schwankungen und zusätzlich in Trend, Saisonkomponente und irreguläre Komponente aufzugliedern. Die Erscheinungsformen und ihre Ursachengruppen werden jedoch meistens nur sehr heuristisch beschrieben.

So versteht man unter dem Trend die allgemeine Entwicklungstendenz einer Zeitreihe, hervorgerufen durch langfristig wirkende Ursachen wie Bevölkerungsentwicklung oder technischen Fortschritt.

Die zyklischen Schwankungen sind wellenförmige Bewegungen um den Trend, die zwar nicht periodisch sein müssen, aber in einigermaßen regelmäßigen Abständen immer wieder auftauchen. Die Ursachen sind weit schwieriger als beim Trend zu beschreiben und in den Wechselwirkungen zwischen bestimmten Verhaltensweisen und den dadurch hervorgerufenen Änderungen des ökonomischen Systems zu suchen.

Die Saisonschwankungen drücken sich in jährlich wiederkehrenden gleichartigen Veränderungen aus. Sie werden hervorgerufen durch Witterungseinflüsse und Kalendereigenschaften.

Die irreguläre Komponente schließlich besteht aus den zufälligen Veränderungen der Zeitreihe, die auf keine der oben genannten Ursachen zurückgeführt werden können.

Die heuristische Beschreibung der Komponente einer Zeitreihe hat einen wesentlichen Mangel, der darin besteht, daß zwar gewisse Ursachengruppen für die Komponenten unterstellt werden, daß aber die Wirkungsweisen, wie eine bestimmte Komponentenform entsteht, nicht erklärt werden. Das liegt einerseits natürlich an der Vielfalt ökonomischer Reihen und der daraus resultierenden Vielfalt entsprechender Mechanismen, die nicht alle in einer allgemeinen Definition untergebracht werden können. Das liegt andererseits aber auch an der generellen Schwierigkeit, für eine bestimmte Zeitreihe alle Ursachengruppen auszumachen, ihre Einzelwirkungen festzustellen, und daraus die Wirkung in ihrer Gesamtheit zu erklären. Man denke nur an die Vielfalt der möglichen Hypothesen für Investitionen und Konsum. Wald (48, S.2) hat diese Problematik unter dem Begriff "äußere Definition der Komponenten einer Zeitreihe" ausführlich erörtert. Für ihn ist die Definition ein Ideal, das zwar nie erreicht werden kann, aber dennoch anzustreben ist. Ein Schritt in diese Richtung, der einer formalen Lösung des Problems teilweise sehr nahekommt, besteht in der zur Zeit Walds noch wenig entwickelten Theorie ökonometrischer Modelle, die wir hier unter diesem Aspekt betrachten wollen.

1.2. Ökonometrische Zeitreihenmodelle

Unter einem solchen Modell versteht man ein System von Beziehungen zwischen den Zeitreihen gewisser Variablen einer Volkswirtschaft. Ein lineares ökonometrisches Modell hat die Form

$$(1.1) \quad A_0 X(t) = A_1 X(t-1) + \dots + A_l X(t-l) + \\ + B_0 Z(t) + \dots + B_r Z(t-r) + U(t)$$

Dabei stellt $X(t) = (X_1(t), X_2(t), \dots, X_p(t))'$ den p-dimensionalen Spaltenvektor der endogenen Variablen, $Z(t) = (Z_1(t), Z_2(t), \dots, Z_q(t))'$, den q-dimensionalen Spaltenvektor der exogenen Variablen und $U(t) = (U_1(t), \dots, U_p(t))'$ wieder einen p-dimensionalen Vektor von latenten Variablen dar.¹⁾

Die Zahl l ist der maximale lag, mit dem endogene Variablen vorkommen, und r der maximale lag der exogenen Variablen. An Annahmen über die auftretenden Größen wollen wir zunächst nur voraussetzen, daß die Matrix A_0 invertierbar sei. Das Modell wird dann vollständig genannt. Unser Ziel ist, den Verlauf der endogenen Variablen innerhalb des Modells durch den Verlauf der exogenen Variablen $Z(t)$, die latenten Variablen $U(t)$ und gewisse Anfangszustände $X(0), \dots, X(l-1)$ zu erklären. Wir bemerken zunächst, daß das Differenzengleichungssystem l-ter Ordnung in $X(t)$ äquivalent ist einem Differenzengleichungssystem erster Ordnung in den Vektoren

$$\begin{bmatrix} X(t) \\ \vdots \\ X(t-l+1) \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} X(t-1) \\ \vdots \\ X(t-l) \end{bmatrix},$$

das ausführlich geschrieben lautet:

$$(1.2) \quad \begin{bmatrix} A & 0 & \dots & 0 \\ 0 & E & & \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & E \end{bmatrix} \cdot \begin{bmatrix} X(t) \\ \vdots \\ X(t-l+1) \end{bmatrix} =$$

1) Die Strukturmatrizen A_j bzw. B_j haben die Dimensionen $p \times p$ bzw. $p \times q$.

$$\begin{aligned}
 &= \begin{bmatrix} A_1 & A_2 & \dots & A_l \\ E & O & \dots & O \\ O & \ddots & \ddots & O \\ O & \dots & O & E & O \end{bmatrix} \cdot \begin{bmatrix} X(t-1) \\ \vdots \\ X(t-1) \end{bmatrix} + \\
 &+ \begin{bmatrix} B_O & B_1 & \dots & B_r \\ O & \dots & \dots & O \\ \vdots & & & \vdots \\ O & \dots & \dots & O \end{bmatrix} \cdot \begin{bmatrix} Z(t) \\ \vdots \\ Z(t-r) \end{bmatrix} + \begin{bmatrix} U(t) \\ O \\ \vdots \\ O \end{bmatrix}
 \end{aligned}$$

Dabei bedeuten die Symbole E und O Einheitsmatrizen und Nullmatrizen von einer solchen Dimension, daß die Multiplikation mit den auftretenden Vektoren sinnvoll wird. Die zusammengesetzte Matrix auf der linken Seite ist mit A_O ebenfalls regulär. Formal kann (1.2) also genau so behandelt werden wie (1.1) wenn $l=1$ und $r=0$ wäre. Wir wollen deshalb die Symbolik gewissermaßen rekursiv benutzen, um (1.2) in geschlossener Form auszudrücken durch

$$(1.3) \quad A_O X(t) = A_1 X(t-1) + B_O Z(t) + U(t).$$

Nach Multiplikation mit A_O^{-1} erhält man daraus die sogenannte reduzierte Form

$$(1.4) \quad X(t) = A X(t-1) + B Z(t) + V(t)$$

mit den Zusammenfassungen

$$A = A_O^{-1} A_1, \quad B = A_O^{-1} B_O, \quad V(t) = A_O^{-1} U(t).$$

Wenn man $X(\tau)$ rekursiv für $\tau = t-1, t-2, \dots, 1$ einsetzt, bekommt man schließlich den Ausdruck

$$(1.5) \quad X(t) = A^t X(0) + \sum_{\tau=0}^{t-1} A^\tau B Z(t-\tau) + \\ + \sum_{\tau=0}^{t-1} A^\tau V(t-\tau)$$

Wir sehen nun, daß sich in einem ökonomischen Modell die endogenen Variablen aus drei Teilen additiv zusammensetzen. Der erste ist endogener Natur. In ihm kommt der Anfangszustand des Systems zum Ausdruck. Dieser Teil stellt die Lösung des homogenen Differenzgleichungssystems (1.4) dar. Der zweite Teil ist vom Verlauf der exogenen Variablen abhängig. Er bildet zusammen mit dem dritten, stochastischen Teil, die singuläre Lösung des Differenzgleichungssystems. Der Verlauf aller drei Anteile hängt außerdem ab von den Strukturmatrizen A_0 , A_1 und B_0 . Insbesondere ist die Matrix $A = A_0^{-1} A_1$ von entscheidender Bedeutung. Hat diese Matrix nur Eigenwerte, die dem Betrage nach kleiner als eins sind, dann konvergiert A^t gegen Null, d.h. in $X(t)$ kommen hauptsächlich die aktuellen Werte $Z(t)$ und $V(t)$ zum Ausdruck, während der Einfluß der Vergangenheitswerte mehr und mehr verschwindet. Wenn allerdings wenigstens einer der Eigenwerte einen Betrag hat, der größer oder gleich 1 ist, kann der Einfluß der Vergangenheitswerte nicht mehr vernachlässigt werden. Dieser instabile Fall ist jedoch nicht wünschenswert und sicherlich auch unwahrscheinlich. Er bedeutet nämlich, daß in der durch das ökonomische Modell beschriebenen Volkswirtschaft die das wirtschaftliche Handeln bestimmenden Größen einen umso höheren Einfluß in der Gegenwart haben, je weiter sie zurück liegen. ~~Da aus diesen Gründen der instabile Fall ganz ausgeschlossen werden kann,~~ wollen wir nicht weiter erörtern.

1.3. Die Komponenten im Rahmen ökonomischer Modelle

Wir wollen uns überlegen, wie die in Abschnitt 1.1. genannten vier Komponenten Trend, Zyklus, Saison und Rest mit den hier genannten endogenen, exogenen und stochastischen Anteilen von (1.5) in Einklang gebracht werden können.

Wir beginnen mit dem Trend, den wir als langfristige Grundrichtung einer Zeitreihe bezeichnet haben. Bei einem großen Teil wirtschaftlich relevanter Größen ist diese Grundrichtung steigend. Aus diesem Grund kann bei einem stabilen Modell nur der exogene Anteil von (1.5) zur Erklärung des Trends herangezogen werden. Die exogenen Variablen $Z(t)$ oder ein Teil davon sind die verursachenden Größen. Die mathematische Form des Verlaufs überträgt sich für eine gewisse Klasse typischer Trendfunktionen langfristig auf die Variablen $X(t)$. Um das einzusehen, nehmen wir an, die $Z(t)$ wären konstant gleich Z , dann ergibt sich für den exogenen Anteil:¹⁾

$$\begin{aligned} \lim_{t \rightarrow \infty} \sum_{\tau=0}^{t-1} A^{\tau} B Z(t-\tau) &= \\ &= \left(\sum_{\tau=0}^{\infty} A^{\tau} \right) B Z = (I-A)^{-1} B Z \end{aligned}$$

Bei Konstanz der exogenen Variablen sind also auch die endogenen Variablen tendenziell konstant. Dieselbe Aussage gilt für die Zuwächse, denn mit den Variablen selbst erfüllen auch die Zuwächse das Differenzengleichungssystem (1.5). Die Überlegungen lassen sich fortsetzen auf die Änderungen der Zuwächse u.s.w.; wenn also $Z(t)$ ein Polynom l-ten Grades ist, dann ist auch $X(t)$ ein Polynom l-ten Grades. Die häufig anzutreffende geometrische Trendentwicklung

¹⁾ Die Existenz der auftretenden Inversen ist wegen der Stabilität des Systems gesichert.

$$(1.6) \quad Z(t) = W^t Z \quad t = 1, 2, \dots$$

mit der reellen Zahl $W > 1$ und einem Anfangszustand Z verdient besondere Beachtung. Besteht die rechte Seite von (1.5) nur aus dem entsprechenden exogenen Anteil, dann erhält man

$$\begin{aligned} X(t) &= W^t \left(\sum_{\tau=0}^{t-1} A^\tau / W^\tau \right) B Z \\ (1.7) \quad &= W^t (I - A^t / W^t) (I - A/W)^{-1} B Z \\ &= W^t Y - A^t Y ; \end{aligned}$$

mit

$$Y = (I - A/W)^{-1} B Z .$$

Bei geometrischem Wachstum der exogenen Variablen besitzen also auch die endogenen Variablen einen dominierenden Term mit der gleichen Wachstumsrate.

Die mit (1.6) angestellte Überlegung ist auch für andere Werte von W richtig., solange W betragsmäßig nicht kleiner als der größte Eigenwert von A ist. Bei stabilen Modellen trifft das insbesondere für $|W| = 1$ zu. Wählen wir nun einmal $W = e^{i\lambda}$ und zum anderen $W = e^{-i\lambda}$ und fassen die Ergebnisse zusammen, dann sieht man, daß aus einer Schwingung

$$\begin{aligned} Z(t) &= \frac{1}{2} e^{i\lambda t} + \frac{1}{2} e^{-i\lambda t} \\ &= Z \cos \lambda t \end{aligned}$$

der exogenen Variablen ein Verlauf

$$X(t) = Y \cos \lambda t - A^t Y$$

der endogenen Variablen wird, in dem die Schwingung mit der gleichen Frequenz dominiert. Interpretieren wir $Z(t)$ als die verursachenden Größen saisonaler Schwankungen, dann können wir daraus den Schluß ziehen, daß die Saisonkomponente ebenfalls im Bereich des exogenen Anteils von (1.5) anzusiedeln ist.

Für die zyklische Komponente ist eine eindeutige Zuordnung auf einen der drei Anteile von (1.5) wesentlich schwieriger. In den bekannten ökonomischen Modellen vom Akzelerator-Multiplikator¹⁾typ z.B. werden die zyklischen Schwankungen in den Schwingungen des homogenen Teils der Differenzengleichung gesehen. Für stabile Modelle bringt diese Interpretation Schwierigkeiten mit sich, denn die in der Realität zu beobachtenden zyklischen Schwankungen haben weder exakte Periodenlängen noch regelmäßig abnehmende Amplituden. Die zyklische Komponente kann deshalb nur sehr beschränkt als endogener Teil von (1.5) angesehen werden. Einen weiteren Erklärungsversuch bieten die Schockhypothesen, wobei die Schocks exogen oder stochastisch angenommen werden können. Es wären dann wieder die Eigenschwingungen des Systems, die immer wieder in unregelmäßigen Abständen neu angeregt werden. Soweit es die Zufallsschocks $V(t)$ betrifft, kann dieser Erklärungsversuch auch anders analysiert werden. Sind nämlich die Komponenten $U_i(t)$ der Vektoren $U(t)$ aus (1.1) Vektoren von unabhängigen Zufallsvariablen, dann sind es auch die $V_i(t)$, und der stochastische Anteil von (1.5) ist für t ein Vektor von gleitenden Durchschnitten. Solche Prozesse sind aber schwach stationär¹⁾ und besitzen, wie die stochastische Theorie zeigen wird, zyklische Eigenschaften, wie sie von den harmonischen Schwingungen her bekannt sind. Die Perioden dieser Schwingungen hängen alleine von der Matrix A ab, sind jedoch nicht

1) Vergleiche Kapitel II

identisch mit den Perioden der Eigenschwingungen des Systems. Zyklische Schwankungen können schließlich auch noch exogen erklärt werden nach dem Muster, wie wir es bei der Saisonkomponente getan haben, sofern unter den exogenen Variablen sich eine befindet, die ein zyklisches Verhalten aufweist.

Als letzte der vier Komponenten einer Zeitreihe bleibt die Restkomponente, die ohne Schwierigkeiten dem stochastischen Teil von (1.5) zugeordnet werden kann.

Wir haben in diesem Abschnitt nach dem Zusammenhang Ursache - Wirkung bei den Komponenten Trend, Zyklus, Saison und Rest einer Zeitreihe gefragt und Antworten für die endogenen Variablen eines ökonomischen Modells gefunden. Das schließt keineswegs aus, daß auch die Komponenten exogener Variablen auf ähnliche Art erklärt werden können. Formal ist dies immer möglich durch "Endogenisierung" dieser Variablen, indem sie auf Zeitfunktionen mit entsprechendem Verlauf zurückgeführt werden. Wir dürfen dabei natürlich nicht übersehen, daß Zeitfunktionen als Erklärung für gewisse Komponentenformen keine Beschreibung der Ursache, sondern nur eine Beschreibung des Verlaufs liefern.

2. Zeitreihenanalyse als Bestimmung der Komponenten

Von der Existenz der vier Komponenten einer Zeitreihe ausgehend, stellt sich die Aufgabe ihrer numerischen Bestimmung. Das methodische Vorgehen hierzu, das schlechthin als Zeitreihenanalyse bezeichnet wird, läßt sich in zwei Schritte einteilen. Der erste besteht in der Festlegung eines Modells, das die Zeitreihe bzw. deren Komponenten in Abhängigkeit von gewissen anderen Zeitreihenverläufen beschreibt. Der zweite Schritt besteht in der Angabe einer Methode zur Berechnung der freien Parameter. Durch das Modell und die wertmäßig festgelegten Parameter lassen sich dann die Komponenten errechnen. Bei manchen Zeitreihenanalyseverfahren wie Trendbereinigungs- oder Saisonbereinigungsverfahren kommt es allerdings nicht auf die Bestimmung aller Komponenten an, sondern auf die separate Bestimmung einer einzelnen.

Es dürfte klar sein, daß der erste Schritt, nämlich die Festlegung eines Modells, der entscheidende ist, der alles weitere mitbeeinflußt. Ein Analyseverfahren ist umso besser zu beurteilen, je näher es der äußeren Definition der Komponenten im Sinne von Wald kommt. In der Regel wird diese Forderung von multivariaten Modellen, die zur Erklärung noch andere ökonomische Reihen heranziehen, eher erfüllt werden als von univariaten Verfahren, die allein von der zu untersuchenden Zeitreihe ausgehen. Die Tatsache, daß dennoch der weitaus größte Teil der Zeitreihenanalyseverfahren univariater Natur ist, bedarf deshalb einer Begründung. Ein Argument ist, daß nicht für jede ökonomische Reihe (exogene Variablen, autonome Größen) das Bedürfnis oder die Möglichkeit zur Erklärung durch andere ökonomische Reihen vorhanden sein wird. Multivariate Modelle benötigen außerdem ein mehrfaches an statistischem Datenmaterial. Wir wollen mit diesen Argumenten allerdings nicht

darüber hinwegtäuschen, daß die fehlende ökonomische Interpretierbarkeit ein Hauptmangel univariater Verfahren ist. Der Mangel wird besonders gravierend bei Verfahren, die abweichend zum oben genannten methodischen Vorgehen gar kein festumrissenes Analysemodell besitzen, sondern nur in der Beschreibung des Rechenverfahrens bzw. einer Schätzmethode bestehen.

Zu den Entscheidungen über das Analysemodell gehört auch die Frage, auf welche Weise die Komponenten zusammenwirken. In den meisten Fällen entscheidet man sich für den additiven Ansatz

$$(2.1) \quad X(t) = D(t) + K(t) + S(t) + U(t) \quad ,$$

wo $X(t)$ die zu untersuchende ökonomische Reihe und die Symbole auf der rechten Seite der Reihenfolge nach die Trend- die zyklische-, die Saison- und die Restkomponente bedeuten. Dieser Ansatz findet auch in der additiven Zusammensetzung der drei Anteile des ökonomischen Modells (1.5) eine gewisse Bestätigung. Allerdings machen die dort angestellten Betrachtungen deutlich, daß insbesondere Trend- und Saisonkomponente unter Umständen in einer anderen Verbindung besser beschrieben werden können, sofern nämlich beide durch die gleiche exogene Größe verursacht werden. Aber auch dieser Fall läßt sich unter (2.1) subsumieren, wenn man $D(t)$ und $S(t)$ geeignet definiert.¹⁾

Die Restkomponente im additiven Ansatz des Analysemodells wird üblicherweise als Zufallsvariable mit Erwartungswert null und von t unabhängiger Varianz angesehen. Die Restwerte sind untereinander unabhängig. Auch die anderen Komponenten können stochastischen Charakter haben. Dazu erinnern wir an die Erklärung der zyklischen Komponente durch den stochastischen Anteil von (1.5) und weisen außerdem darauf hin, daß die exogenen Variablen in einem ökonomischen Modell durch solche wurden ja Trend und Saison erklärt - ebenfalls Zufallsvariablen sein können.

1) Vergleiche dazu: Wald (48, S.3 f)

3. Zwei methodische Beispiele

3.1. Ein Regressionsmodell

Bei diesem Modell sind für die Komponenten $D(t)$, $K(t)$ und $S(t)$ gewisse Funktionsklassen vorgegeben, z.B. Polynome $d_1(t)$ für den Trend, Sinus- und Cosinusschwingungen $k_1(t)$ mit bestimmten Perioden für die zyklische Komponente und periodische Funktionen $s_1(t)$ mit der Periode ein Jahr für die Saisonkomponente. Die Komponenten sind Linearkombinationen dieser Ansatzfunktionen, d.h.:

$$\begin{aligned} D(t) &= a_1 d_1(t) + a_2 d_2(t) + \dots + a_p d_p(t) \\ (3.1) \quad K(t) &= b_1 k_1(t) + b_2 k_2(t) + \dots + b_q k_q(t) \\ S(t) &= c_1 s_1(t) + c_2 s_2(t) + \dots + c_r s_r(t) \end{aligned}$$

mit $m = p + q + r$ unbekannten Parametern a_i , b_i , c_i . Die irreguläre Komponente $U(t)$ sei eine Folge von Zufallsvariablen mit den Eigenschaften

$$\begin{aligned} E(U(t)) &= 0, \quad \text{Var}(U(t)) = \sigma^2, \\ (3.2) \quad \text{Cov}(U(t)U(s)) &= 0 \quad \text{für } t \neq s. \end{aligned}$$

Durch Zusammenfassung gemäß (2.1) erhalten wir den Regressionsansatz

$$X(t) = a_1 d_1(t) + \dots + c_r s_r(t) + U(t).$$

Unter der Voraussetzung der linearen Unabhängigkeit der Ansatzfunktionen lassen sich bekanntermaßen die Koeffizienten mit Hilfe der Methode der kleinsten Quadrate als beste, lineare, unverzerrte

Schätzungen bestimmen. Fassen wir die hier auftretenden Größen zu den folgenden Vektoren und Matrizen

$$X = \begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(n) \end{bmatrix} \quad U = \begin{bmatrix} U(1) \\ U(2) \\ \vdots \\ U(n) \end{bmatrix}$$

$$\beta = \begin{bmatrix} a_1 \\ \vdots \\ b_i \\ \vdots \\ c_r \end{bmatrix} \quad A = \begin{bmatrix} d_1(1) & \dots & k_i(1) & \dots & s_r(1) \\ d_1(2) & & & & \vdots \\ \vdots & & & & \vdots \\ d_1(n) & \dots & \dots & \dots & s_r(n) \end{bmatrix}$$

zusammen, dann erhält man das Regressionsmodell und die Schätzung in der Darstellung

$$(3.3) \quad X = A \beta + U; \quad \hat{\beta} = (A' A)^{-1} A' X.$$

Die Schätzwerte $\hat{a}_i$, $\hat{b}_i$ und $\hat{c}_i$ des Vektors $\hat{\beta}$ in (3.1) eingesetzt, ergeben die Schätzungen $\hat{D}(t)$, $\hat{K}(t)$ und $\hat{S}(t)$ für die Komponenten. Schließlich erhält man für den Vektor der Erwartungswerte $E(X(t))$, die man auch als systematische Komponente bezeichnet, eine Schätzung

$$(3.4) \quad \hat{X} = A \hat{\beta}$$

Über die aus der Regressionstheorie bekannten Schätzvarianzen der Parameter lassen sich dann Wahrscheinlichkeitsaussagen über

den tatsächlichen Verlauf der Komponenten herleiten.

3.2. Die Methode der gleitenden Durchschnitte

Das gerade betrachtete Regressionsmodell baut insbesondere auf der explizit angegebenen Hypothese auf, daß die Komponenten zu bestimmten Funktionsklassen gehören. Dies setzt a priori Entscheidungen voraus, die man mit der Vorgehensweise nach der Methode der gleitenden Durchschnitte zu umgehen versucht. Hierbei werden Trend und zyklische Komponente zu der glatten Komponente $G(t)$ zusammengefaßt. Der Ansatz lautet dann

$$X(t) = G(t) + S(t) + U(t) \quad .$$

Handelt es sich um Vierteljahresdaten, dann wird eine Schätzung der glatten Komponente, z.B. als gleitender Durchschnitt

$$\hat{G}(t) = \frac{1}{4} \left(\frac{1}{2} X(t-2) + X(t-1) + X(t) + X(t+1) + \frac{1}{2} X(t+2) \right)$$

bestimmt. Wenn wir die hier auftretenden Gewichte mit W_j , $j = -2, -1, \dots, 2$ bezeichnen, können wir schreiben:

$$(3.5) \quad \hat{G}(t) = \sum_j W_j X(t+j) = \sum_j W_j G(t+j) + \sum_j W_j S(t+j) + \\ + \sum_j W_j U(t+j).$$

Hierbei können nun zwei wünschenswerte Effekte auftreten.

a) Falls die Saisonkomponente die Abweichung vom Jahresdurchschnitt ergibt und streng periodisch ist, gilt

$$\sum_j W_j S(t+j) = 0 \quad .$$

Somit erhält man als Schätzung:

$$\hat{G}(t) = \sum_j W_j G(t+j) + \sum_j W_j U(t+j).$$

Es handelt sich also um eine Saisonbereinigung.

b) Falls die irreguläre Komponente die Bedingung (3.2) erfüllt, wird

$$\begin{aligned} \text{Var} \left(\sum_j W_j U(t+j) \right) &= \sum_j \text{Var} (W_j U(t+j)) \\ &= \sum_j W_j^2 \sigma^2 < \sigma^2 \end{aligned}$$

d.h. die Varianz der irregulären Komponente von $\hat{G}(t)$ hat sich verringert. Für unser Beispiel gilt:

$$\sum_j W_j^2 = 14/64.$$

Trotzdem liefert $\hat{G}(t)$ im allgemeinen eine verzerrte Schätzung der glatten Komponente, denn nur wenn diese linear verläuft gilt:

$$E(\hat{G}(t)) = \sum_j W_j G(t+j) = G(t) \text{ für alle } t,$$

wie man leicht nachrechnet.

Wir stellen also fest, daß die zunächst variabler scheinende Methode der gleitenden Durchschnitte erst nach sehr einschränkenden Voraussetzungen, die man gerade umgehen wollte, vernünftige Schätzergebnisse liefert.

4. Die Beurteilung eines Analyseverfahrens

4.1. Kriterien für das Analysemodell

Wir untersuchen in diesem und dem nächsten Abschnitt die Frage, wie man ein Verfahren zur Zeitreihenanalyse beurteilen kann. Wir unterteilen dabei wieder in Analysemodell und Rechen- bzw. Schätzmethode zur Bestimmung der freien Parameter. Wenn man Schumpeters Meinung (41, S. 166 f.) folgt "if trend analysis is to have any meaning, it can derive it only from previous theoretical considerations, which must not only guide us in interpreting results, but also in choosing the method...", dann dürfte klar sein, daß in erster Linie das Analysemodell als Beurteilungsgrundlage in Frage kommt. Wir sahen allerdings schon im letzten Abschnitt, daß univariate Analyseverfahren eine ökonomisch inhaltliche Beurteilung nur in sehr beschränktem Maße zulassen. Es kann sich nur noch darum handeln zu vergleichen, ob die im Modell postulierten Eigenschaften der einzelnen Komponenten mit den aus der ökonomischen Theorie zu erwartenden Eigenschaften übereinstimmen. Insbesondere ist festzustellen, ob das Analysemodell die für die zu untersuchende Reihe üblichen Komponentenformen impliziert.

Die Beurteilung verlagert sich also vom Gehalt mehr in Richtung zur Diskription. Das hat die Folge, daß auch die Frage interessant wird, ob die Beschreibung mit wenigen und dafür informations-trächtigen oder mit relativ vielen informationsarmen Größen geschieht, wobei der erste Fall dem zweiten vorzuziehen ist. Box und Jenkins (6, S. 17) nennen dies das Prinzip der Sparsamkeit. Wir betonen nochmals, daß dieses Prinzip erst durch mangelnden ökonomischen Gehalt im Modell an Bedeutung gewinnt.

Ein weiteres Beurteilungskriterium eines Zeitreihenmodells ist seine Prognosefähigkeit. Ein Modell muß man dann positiv beurteilen, wenn seine Schätzungen zur Erstellung von Prognosen und Wahrscheinlichkeitsaussagen über den Prognosefehler benutzt werden können, ohne daß man den Rahmen des Modells verlassen muß.

Wir wollen die drei Beurteilungskriterien an den beiden Beispielen überprüfen. Da beim Regressionsmodell von Fall zu Fall die Ansatzfunktionen geeignet gewählt werden können, ist die Frage nach den richtigen Eigenschaften der Komponente positiv zu beantworten. Auch das Prinzip der Sparsamkeit ist erfüllt, denn die Information, die in den Zeitreihenwerten $X(t)$, $t=1, \dots, n$ enthalten ist, konzentriert sich - große n vorausgesetzt - auf die relativ wenigen

Parameter a_i , b_i und c_i . Schließlich eignet sich das Regressionsmodell auch zur Berechnung von Prognosen, die sich über die geschätzten Parameter und die Werte der Ansatzfunktionen im zu prognostizierenden Zeitpunkt auf einfache Weise anbieten. Bezeichnen wir nämlich die Prognose zum Zeitpunkt n für den Zeitpunkt $n + \tau$ mit $X(n, \tau)$ und mit $a(n + \tau)$ den Spaltenvektor $(d_1(n + \tau), \dots, s_r(n + \tau))'$ der Ansatzfunktionen zum Zeitpunkt $n + \tau$, dann erhält man als Prognose :

$$X(n, \tau) = a'(n + \tau) \hat{\beta}$$

mit dem Prognosefehler

$$X(n + \tau) - X(n, \tau) = U(n + \tau) - a'(n + \tau) (\beta - \hat{\beta}).$$

Da $U(n + \tau)$ für $\tau > 0$ unabhängig ist von der Schätzung $\hat{\beta}$, die ja auf den Werten $U(1)$, $U(2)$, $\dots$, $U(m)$ beruht, errechnet man für die Varianz des Prognosefehlers

$$\text{Var} (X(n+\tau) - X(n, \tau)) = \sigma^2 - a'(n+\tau) \text{Var}(\beta) a(n+\tau)$$

Aus diesem Ausdruck lassen sich Wahrscheinlichkeitsaussagen für das Eintreffen der Prognose ableiten.

Bei der Methode der gleitenden Durchschnitte entfällt der erste Beurteilungspunkt, da die Eigenschaften der Komponente nicht in einem explizit angegebenen Analysemodell, sondern nur implizit durch die Rechenmethode gegeben sind. Das Prinzip der Sparsamkeit ist nicht erfüllt, denn die geschätzte glatte Komponente muß durch genau so viele Einzelwerte angegeben werden, wie Ausgangswerte vorhanden sind. Eine Informationskonzentration hat nicht stattgefunden. Dabei muß außerdem noch die Schwierigkeit berücksichtigt werden, an den Rändern der Zeitreihe überhaupt Schätzwerte zu bekommen. Für Prognosewerte ist die Methode ebenfalls nicht geeignet. Wie die Randwerte kann man Prognosewerte erst durch zusätzliche Ergänzungen bzw. Erweiterungen der Methode erhalten.

4.2. Forderungen an die Schätzmethode

Grundsätzlich wollen wir fordern, daß die Schätzmethode konsistente Schätzungen für die Komponenten liefert, d.h. für $n \rightarrow \infty$ sollen die Schätzwerte im quadratischen Mittel gegen die wahren Komponentenwerte - womit die im Modell postulierten gemeint sind - konvergieren. Diese Forderung ist von gegebenen Eigenschaften des Modells im allgemeinen nicht zu trennen. Wir haben es also mit einer Forderung an das Analyseverfahren schlechthin zu tun. Das betrachtete Regressionsverfahren zur Zeitreihenanalyse genügt dieser Forderung, die Methode der gleitenden Durchschnitte jedoch nicht, wie wir bereits gesehen haben.

Lovell (31) hat eine Reihe weiterer Bedingungen untersucht, die wir uns an den beiden methodischen Beispielen klarmachen wollen. Mit $L X(t)$ bezeichnen wir dabei irgendeine aus der Zeitreihe $X(t)$ gewonnene Schätzung der Komponenten eines Verfahrens.

Die erste Bedingung ist die Summenerhaltung. Sie lautet

$$(4.1) \quad L(X(t) + Y(t)) = L X(t) + L Y(t)$$

und ist für beide hier genannten Beispiele erfüllt, da es sich bei (3.4) und (3.5) um lineare Schätzungen handelt.

Analog zu (4.1) erfüllt ein produkterhaltendes Verfahren die Bedingung

$$L(X(t) Y(t)) = L X(t) L Y(t)$$

Das bedeutet, daß zum Beispiel die geschätzte Komponente einer Nominalreihe tatsächlich auch als Produkt der geschätzten Komponenten von Preis und Realwert gewonnen werden kann. Die Forderung ist allerdings nur sinnvoll bei einem multiplikativen Ansatz für die Komponenten.

Ein Verfahren ist idempotent, wenn seine Anwendung auf die geschätzte Reihe das gleiche Ergebnis bringt, d.h.

$$L(L X(t)) = L X(t)$$

Das Regressionsverfahren genügt dieser Bedingung, wie man nach (3.3) und (3.4) an der Beziehung

$$(A'A)^{-1} A' \hat{X} = (A'A)^{-1} A'A \hat{\beta} = \hat{\beta}$$

abliest. Für die Methode der gleitenden Durchschnitte gilt dies

nicht , da die ihre Anwendung auf die $\hat{G}(t)$ nur dann zum gleichen Ergebnis führt, wenn diese linear verlaufen.

Weiter wird die Orthogonalität angeführt, die durch

$$(4.2) \quad \sum_{t=1}^n (X(t) - L X(t)) L X(t) = 0$$

beschrieben wird. Diese Beziehung ergibt sich beim Regressionsverfahren aus den Normalgleichungen. Bei den gleitenden Durchschnitten kann (4.2) nur für Spezialfälle nachgewiesen werden. Während Idem-potenz so interpretiert werden kann, daß die geschätzte Komponente keine Reste anderer Komponenten mehr enthält, bedeutet die Orthogonalität umgekehrt , daß auch die Reste keine Anteile mehr von der geschätzten Komponente enthalten.

Die fünfte und letzte der von Lovell genannten Eigenschaften ist die Symmetrie

$$\frac{\partial L X(t)}{\partial X(t+s)} = \frac{\partial L X(t+s)}{\partial X(t)}$$

eine Eigenschaft, die wiederum beide Verfahren besitzen. Zum Nachweis für das Regressionsverfahren betrachte man (3.4), woraus sich durch Einsetzen von $\hat{\beta}$ aus (3.5) und Differenzieren die Matrix der partiellen Ableitungen

$$\frac{\partial \hat{X}}{\partial X} = A(A'A)^{-1} A'$$

ergibt, die symmetrisch ist. Für den gleitenden Durchschnitt erhält man unter Benutzung von (3.5)

$$\frac{\partial \hat{X}(t)}{\partial X(t+s)} = W_s \quad \frac{\partial \hat{X}(t+s)}{\partial X(t)} = W_{-s}$$

Die Durchschnittsbildung ist also dann symmetrisch, wenn wie in unserem Beispiel das Gewichtsschema symmetrisch ist.

Lovell hat in (31, S. 1005) gezeigt, daß es kein sinnvolles Verfahren gibt, das alle fünf Eigenschaften erfüllen kann.

Nur die trivialen Verfahren

$$L X(t) = X(t) \quad \text{und} \quad L X(t) = 0$$

leisten das. Im Rahmen der hier betrachteten Kriterien ist also das Regressionsverfahren optimal, da nur die Produkterhaltung verletzt wird. Bei dieser Bewertung ist noch zu berücksichtigen, daß jedes summenerhaltende Verfahren, das zwei der drei Eigenschaften Idempotenz, Orthogonalität und Symmetrie erfüllt, auch automatisch die dritte besitzt.

II. Wahrscheinlichkeitstheoretische Grundlagen stationärer Prozesse

5. Stochastische Prozesse

Ein stochastischer Prozeß soll als wahrscheinlichkeitstheoretisches Modell dazu dienen, den Zusammenhang der beobachteten Werte einer Zeitreihe zu erklären. Geht man davon aus, daß jeder Wert $x(t)$ die Realisation einer Zufallsvariablen $X(t)$ ist, dann liegt es nahe, einen stochastischen Prozeß, den wir kurz mit X bezeichnen wollen, durch eine Familie von Zufallsvariablen $X(t)$ zu definieren, die von einem gewissen Parameter t abhängt, der in einer Parametermenge T variiert, d.h.:

$$X = (X(t)/t \in T).$$

Diese formale Definition ist sehr umfassend und trifft beispielsweise auch auf einfache Stichproben zu. Die Bezeichnung "Prozeß" deutet jedoch darauf hin, daß in den Anwendungen die Variablen $X(t)$ des Prozesses nicht unabhängig voneinander sind und ihr Aufeinanderfolgen eine wesentliche Rolle spielt. Erst dadurch wird die Theorie der stochastischen Prozesse zu einer stochastischen Theorie der Zeitreihen. Die Parametermenge T wird deshalb als "Zeit" interpretiert und durch eine im allgemeinen unendliche Teilmenge der reellen Zahlengeraden dargestellt. Man spricht von einem diskreten Prozeß, wenn T abzählbar ist und von einem kontinuierlichen Prozeß, wenn t alle Werte eines Intervalls annehmen kann.

Was wir als eine Zeitreihe beobachten, ist eine Realisation des stochastischen Prozesses X , die wir als eine Menge

$$x = (x(t)/t \in T)$$

von Realisationen der Prozeßvariablen $X(t)$, $t \in T$, erklären. Jedes x ist demnach ein Punkt in einem Stichprobenraum E , dessen Dimension gleich der Kardinalzahl von T ist, und dessen Koordinatenräume die Menge der reellen Zahlen sind. Andererseits kann E auch als die Menge aller auf T erklärten reellen Funktionen aufgefaßt werden. Deshalb wird x auch zufällige Funktion genannt.

Die Definition des Prozesses als eine Familie von Zufallsvariablen beinhaltet, daß jeweils endlich viele $X(t_i)$, $i = 1, 2, \dots, m$, deren Parameter wir zur Menge $M \subset T$ zusammenfassen, eine gemeinsame Verteilungsfunktion

$$(5.1) \quad F_M(x_1, x_2, \dots, x_m) = P(X(t_i) \leq x_i, t_i \in M)$$

besitzen. Durch die Menge aller dieser Verteilungsfunktionen - d.h. wenn M über alle endlichen Teilmengen von T variiert - wird auf E eine Wahrscheinlichkeitsverteilung induziert, die als Verteilung des Prozesses X bezeichnet wird. Die Borelmengen in E sind diejenigen, die durch die üblichen Mengenoperationen aus den Mengen $A \subset E$ von der Form

$$A = \{x \in E / X(t_i) = x_i, t_i \in M\}$$

erzeugt werden. Das Wahrscheinlichkeitsverhalten des Prozesses wird also ausgehend von dem Wahrscheinlichkeitsverhalten in endlich vielen Zeitpunkten beschrieben.

Aus dem oben geschilderten ersieht man, daß die Definition eines stochastischen Prozesses letztlich darin besteht, eine geeignete Wahrscheinlichkeitsverteilung auf dem Raum E aller möglichen zufälligen Funktionen zu finden. Man kann dieses Ziel auch direkt angehen, indem man sich für jede endliche Teilmenge M aus T eine m -dimensionale Verteilungsfunktion $F_M(x_1, x_2, \dots, x_m)$

vorgibt. Kolmogorov (26) hat gezeigt, daß diese Verteilungsfunktionen für verschiedene Indexmengen M und N beliebig sein können, bis auf eine Verträglichkeitsforderung für den Fall $M \subset N$:

$$(5.2) \quad F_N(x_1, \dots, x_m, x_{m+1}, \dots, x_n) \rightarrow F_M(x_1, x_2, \dots, x_m)$$

$$\text{für } x_{m+1}, \dots, x_n \rightarrow \infty.$$

Dadurch wird sichergestellt, daß sich die verschiedenen Verteilungen auch als Randverteilungen untereinander richtig verhalten. Ist die Forderung (5.2) erfüllt, dann erhält man, indem man (5.1) als Definition für die rechte Seite $P(\dots)$ auffaßt, wie oben beschrieben die Verteilung eines stochastischen Prozesses X auf E. Die zuletzt betrachtete Definition eines stochastischen Prozesses unterscheidet sich von der ersten in der Tatsache, daß dort die Beziehungen (5.1) und (5.2) automatisch über den Begriff "Familie von Zufallsvariablen" mitgeliefert werden.

Die Eigenschaften der Verteilungsfunktionen (5.1) dienen natürlicherweise zur Klassifikation stochastischer Prozesse. Wie oben schon erwähnt, spielt dabei insbesondere die stochastische Abhängigkeit der Prozeßvariablen eine Rolle, die u.a. in den Kovarianzen $\text{Cov}(X(t), X(s))$ oder auch in den entsprechenden nichtzentralen Momenten $E(X(t) \cdot X(s))$ für t, s aus T zum Ausdruck kommt. Da der Zusammenhang

$$(5.3) \quad \text{Cov}(X(t), X(s)) = E(X(t) X(s)) - E(X(t)) E(X(s))$$

besteht, können beide Größen gleichermaßen zur Beschreibung verwandt werden. Die Menge der Kovarianzen in Abhängigkeit von t und s heißt auch Kovarianzfunktion oder, wenn es darauf ankommt, von den Kovarianzen zwischen zwei Variablen

verschiedener Prozesse zu unterscheiden, Autokovarianzfunktion. Für $t = s$ ist der Wert der Kovarianzfunktion per def. gleich der Varianz von $X(t)$.

Bisher sind wir davon ausgegangen, daß die $X(t)$ reellwertige Zufallsvariablen sind. Ein stochastischer Prozess $Z = (Z(t))/t \in T$ heißt komplexwertig, wenn die Zufallsvariablen $Z(t)$ komplexe Werte annehmen können. Jede komplexe Zufallsvariable läßt sich zerlegen in

$$Z(t) = U(t) + iV(t), \quad t \in T,$$

wobei der Realteil $U(t)$ und der Imaginärteil $V(t)$ reelle Zufallsvariablen sind. Entsprechend sind $U = (U(t))/t \in T$ und $V = (V(t))/t \in T$ reelle stochastische Prozesse, und es gilt $Z = U + iV$. Der Prozeß ist äquivalent dem zweidimensionalen reellen Prozeß $Z = (U, V)$; die Verteilung von Z ist deshalb bestimmt durch die gemeinsame Verteilung der Prozesse U und V , d.h. durch alle gemeinsamen Verteilungsfunktionen der Form (5.1) mit je endlich vielen $U(t)$ und endlich vielen $V(t)$. Als Kovarianzfunktion eines komplexwertigen Prozesses definieren wir

$$(5.4) \quad \text{Cov} (Z(t), Z(s)) = E (Z(t) \overline{Z(s)}) - E(Z(t)) \overline{E(Z(s))};$$

dabei ist $\overline{Z(s)} = U(s) - iV(s)$ die konjugiert komplexe Zufallsvariable von $Z(s)$. Im reellen Fall stimmt (5.4) mit (5.3) überein. Durch Einsetzen von $Z(t)$ und $\overline{Z(s)}$ in die rechte Seite von (5.4) rechnet man leicht nach, daß allgemein

$$\begin{aligned} \text{Cov} (Z(t), Z(s)) &= \text{Cov} (U(t), U(s)) + \text{Cov} (V(t), V(s)) + \\ &\quad + i\text{Cov} (V(t), U(s)) - i\text{Cov} (U(t), V(s)) \end{aligned}$$

gilt. Die Kovarianzfunktion im komplexen Fall kann also auf die Autokovarianzen der Prozesse U und V und deren gegenseitigen sogenannten Kreuzkovarianzen zurückgeführt werden. Wir notieren noch einige evidente Eigenschaften:

$$E(Z(t)\overline{Z(s)}) = \overline{E(Z(s)\overline{Z(t)})}$$

$$(5.5) \quad \text{Cov}(Z(t), Z(s)) = \overline{\text{Cov}(Z(s), Z(t))}$$

$$E(|Z(t)|^2) = E(U^2(t)) + E(V^2(t))$$

$$\text{Var}(Z(t)) = \text{Var}(U(t)) + \text{Var}(V(t))$$

Komplexe Prozesse sind formal von Bedeutung, um insbesondere die Theorie der stationären Prozesse eleganter formulieren zu können. Wir werden im folgenden davon gelegentlich Gebrauch machen. Wenn jedoch nicht ausdrücklich etwas anderes gesagt wird, verstehen wir unter $X(t)$, $t \in T$ einen reellen Prozeß, der das stochastische Modell für eine in der Praxis beobachtete Zeitreihe darstellt.

6. Stationäre stochastische Prozesse

Wir betrachten einen stochastischen Prozeß X in der Zeit T . Es sei M eine endliche Parametermenge aus T und $M + d$ die Menge der um die Zeitdifferenz d verschobenen Parameter $t_1 + d$, $t_1 \in M$. Der Prozeß X wird strikt stationär oder stationär im engeren Sinne genannt, wenn für beliebige Zeitdifferenzen mit M und $M + d$ aus T die Verteilungsfunktionen (5.1) die Bedingung

$$(6.1) \quad F_M(x_1, \dots, x_m) = F_{M+d}(x_1, \dots, x_m)$$

erfüllen. Die gemeinsame Verteilung von endlich vielen Variablen des Prozesses hängt also nur vom zeitlichen Abstand untereinander ab und ist unabhängig von einer Zeitverschiebung. Betrachtet man nur die eindimensionalen und zweidimensionalen Verteilungen, dann ergeben sich aus (6.1) folgende wesentliche Eigenschaften:

$$(6.2) \quad E(X(t)) = \text{const.}$$

d.h. alle Prozeßvariablen haben den gleichen Erwartungswert,

$$(6.3) \quad \text{Cov}(X(t), X(s)) = \text{Cov}(X(t-s), X(0))$$

d.h. die Kovarianzfunktion ist eine Funktion, die nur vom zeitlichen Abstand $\tau = t-s$ abhängt. Benutzt man anstelle der Kovarianzen die nichtzentralen Momente, dann gilt auch

$$(6.4) \quad E(X(t) X(s)) = E(X(t-s) X(0)) = c(t-s) .$$

Die Gleichungen (6.2) - (6.4) dienen ihrerseits zur Definition von gewissen Verallgemeinerungen. So wird ein Prozeß mittelwertstationär genannt, wenn die Gleichung (6.2) gilt. Er wird kovarianzstationär genannt, wenn (6.4) erfüllt ist. Beide Begriffe widersprechen jedoch einer heuristischen Vorstellung von Stationarität, da im ersten Fall z.B. die Varianzen und im zweiten Fall die Erwartungswerte noch unbeschränkt wachsen können. Dies ist nicht mehr möglich bei den sogenannten schwach stationären Prozessen, die auch stationär im weiteren Sinne genannt werden. Hierfür gibt es zwei Definitionen.

Definition 1: Ein stochastischer Prozeß heißt schwach stationär, wenn (6.2) und (6.3) erfüllt ist.

Definition 2: Ein stochastischer Prozeß heißt schwach stationär, wenn (6.4) erfüllt ist.

Mit Hilfe von Gleichung (5.3) findet man leicht, daß aus (6.2) und (6.3) die Beziehung (6.4) folgt, sowie aus (6.2) und (6.4) die Beziehung (6.3). Die Definition 2 umfaßt also alle Prozesse, die in Definition 1 enthalten sind. Äquivalenz besteht aber nur, wenn $E(X_t) = \text{const.}$ gilt. Jeder strikt stationäre Prozeß ist natürlich auch schwach stationär. Die Umkehrung dieser Aussage gilt i.a. jedoch nicht. Sie trifft höchstens für Spezialfälle zu, z.B. für den Gaußschen Prozeß (vgl. Abschnitt 5).

Soweit überschaubar, benutzen alle Autoren, die schwach stationäre Prozesse zur Untersuchung ökonomischer Reihen heranziehen, die Definition 1. Mit Ausnahme von Doob gilt das auch für jene Autoren, die sich mit der stochastischen Theorie alleine befassen. Doob stellt jedoch klar heraus, daß die wesentlichen Eigenschaften schwach stationärer Prozesse - gemeint ist insbesondere die Spektraldarstellung - alleine aus Definition 2 hervorgehen. "Usually the added condition that $E(X(s))$ does not depend on s is imposed. This condition is unnatural mathematically and has nothing to do with the essential properties of interest in these processes, ..." (7, S. 95).

Von daher gesehen ist die übliche Definition 1 unnötigerweise einschränkend. Sie postuliert Konstanz der Erwartungswerte, obwohl nach Doobs Definition noch eine gewisse Variabilität möglich ist. Diese kommt zum Ausdruck, wenn man die Beziehung

$$c(|t-s|) = \text{Cov}(X(t), X(s)) + E(X(t)) E(X(s))$$

beachtet. Für $t = s$ erhält man z.B.

$$c(0) = \text{Var}(X(t)) + E^2(X(t)).$$

Erwartungswerte, Varianzen und Kovarianzen können sich also noch in einem durch die Funktion $c(0)$ festgelegten Bereich verändern.

Wir wollen aus diesen Gründen schwache Stationarität mit Definition 2 einführen. Wie Doob bezeichnen wir die Funktion $c(|t-s|) = E(X(t) X(s))$ als Kovarianzfunktion, obwohl es sich tatsächlich erst dann um Kovarianzen handelt, wenn $E(X(t)) = 0$ ist für alle t und damit wie oben bemerkt, auch Definition 1 zutrifft. Wir sprechen von einem stationären Prozess schlechthin, wenn es auf die Unterscheidung strikt oder schwach nicht ankommt.

Ist X ein komplexwertiger stationärer Prozeß, dann wollen wir ihn in Erweiterung des reellen Falles schwach stationär nennen, wenn die Kovarianzfunktion $E(X(t)\overline{X(s)})$ ebenfalls nur von $t-s$ abhängt, d.h.:

$$E(X(t) \overline{X(s)}) = E(X(t-s) \overline{X(0)}) = c(t-s) .$$

Wie man aus (5.5) abliest, hat $c(\cdot)$ die Eigenschaft

$$c(t-s) = \overline{c(s-t)}$$

Erst im reellen Fall, wenn die Kovarianzfunktion selber reell wird, wird daraus die bekannte Gleichheit $c(s-t) = c(t-s)$.

7. Spektraldarstellungen bei stationären Prozessen

7.1. Prozesse mit orthogonalen Zuwächsen

Diese Prozesse sind eine Verallgemeinerung der Prozesse mit unabhängigen Zuwächsen, die unter anderem als Modell für die Brownsche Molekularbewegung von Wiener in (52) untersucht worden sind. Der sogenannte Wienerprozeß wird auch von Doob (7, S. 216) behandelt. Zur Illustration wollen wir uns dieses Modell kurz vor Augen führen.

Die Zufallsvariable $X(t)$ stehe dabei für eine kartesische Koordinate eines Teilchens in einer Gasmenge, das den unkontrollierbaren Zusammenstößen mit den übrigen Molekülen des Gases ausgeliefert ist. Das Teilchen werde im Zeitintervall $a \leq t \leq b$ beobachtet und befinde sich am Anfang o.B.d.A. im Punkt $X(a) = 0$. Da die Zusammenstöße nach allen Richtungen gleichmäßig erfolgen, kann man annehmen, daß die Zuwächse - in diesem Fall die Ortsveränderungen zwischen zwei Beobachtungszeitpunkten - $X(t) - X(s)$ den Erwartungswert null haben, und daß zwei Ortsveränderungen, die aufeinanderfolgen, unabhängig sind. Es gilt dann

$$E((X(t_2) - X(s_2)) (X(t_1) - X(s_1))) = 0$$

(7.1)

$$\text{für } s_1 < t_1 \leq s_2 < t_2.$$

Infolge dieser Eigenschaften hat der Ort $X(t)$ des Teilchens als Summe aller bis zum Zeitpunkt t stattgefundenen Veränderungen ebenfalls den Erwartungswert null, aber eine mit t sich ständig vergrößernde Varianz. Darin kommt die Tatsache zum Ausdruck, daß mit fortschreitender Zeit die Entfernung des Teilchens vom Nullpunkt immer ungenauer vorausgesagt werden

kann.

Neben der hier geschilderten praktischen Anwendungsmöglichkeit haben Prozesse mit orthogonalen Zuwächsen, die im reellen Fall durch (7.1) definiert werden, auch eine formale Bedeutung. Wir betrachten dazu einen komplexen Prozeß Z mit stetiger Parametermenge $a \leq t \leq b$. Der Prozeß möge orthogonale Zuwächse besitzen. Das bedeutet im komplexen Fall

$$(7.2) \quad E((Z(t_2) - Z(s_2)) \overline{(Z(t_1) - Z(s_1))}) = 0$$

für $s_1 < t_1 \leq s_2 < t_2$.

In symbolischer Form lautet (7.2) auch

$$E(dZ(t) \overline{dZ(s)}) = 0 \quad \text{für} \quad t \neq s.$$

Für $a \leq t \leq b$ definieren wir eine Funktion

$$(7.3) \quad F(t) = E(|Z(t) - Z(a)|^2),$$

die wir als Varianzfunktion des orthogonalen Prozesses Z bezeichnen wollen. Analog zu den Betrachtungen beim Wienerprozeß muß $F(t)$ eine monoton steigende Funktion sein. Aufgrund der Orthogonalität findet man für die Zuwächse von $F(t)$:

$$(7.4) \quad F(t) - F(s) = E(|Z(t) - Z(s)|^2)$$

$$\text{oder} \quad dF(t) = E(|dZ(t)|^2).$$

Setzt man wie üblich voraus, daß die Varianzen von Z beschränkt sind, dann hat $F(t)$ bis auf eine Normierung alle Eigenschaften einer Verteilungsfunktion in $a \leq t \leq b$.

Sie eignet sich deshalb zur Bildung eines Stieltjesintegrals

$$(7.5) \quad c(k) = \int_a^b k(t) dF(t)$$

mit einer gewissen Funktion $k(t)$. Ist diese stetig, dann erhält man (7.5) als Grenzwert von Riemann-Stieltjessummen, d.h.

$$c(k) = \lim_{n \rightarrow \infty} \sum_{i=1}^n k(t_i') (F(t_i) - F(t_{i-1})) .$$

Dabei ist $a = t_0 < t_1 < \dots < t_n = b$ eine Intervallzerlegung mit $t_i - t_{i-1} \rightarrow 0$ für $n \rightarrow \infty$ und $t_{i-1} \leq t_i' \leq t_i$.

Analog zu (7.5) kann man bezüglich Z auch sogenannte stochastische Integrale

$$(7.6) \quad X(h) = \int_a^b h(t) dZ(t)$$

definieren (7, S. 246). Bei stetigem $h(\cdot)$ handelt es sich hierbei um den Grenzwert im quadratischen Mittel (limes in the mean)¹⁾

$$(7.7) \quad X(h) = \text{l.i.m.}_{n \rightarrow \infty} \sum_{i=1}^n h(t_i') (Z(t_i) - Z(t_{i-1})) .$$

Wegen der Orthogonalität des Prozesses Z erhält man für die Varianz der rechts stehenden Summe

$$\begin{aligned} (7.8) \quad E(\Sigma \cdot \overline{\Sigma}) &= \sum_{i=1}^n |h(t_i')|^2 E(|Z(t_i) - Z(t_{i-1})|^2) \\ &= \sum_{i=1}^n |h(t_i')|^2 (F(t_i) - F(t_{i-1})) . \end{aligned}$$

1) Zur Definition des Grenzwertes vergl. (7, S. 8).

Die Existenz der Zufallsvariablen $X(h)$ - um eine solche handelt es sich bei dem stochastischen Integral (7.5) - ist sichergestellt, wenn (7.8) konvergiert, d.h. wenn $c(|h|^2)$ nach (7.5) existiert.

Die Varianzfunktion $F(t)$ kann in drei Komponenten zerlegt werden

$$(7.9) \quad F(t) = F_1(t) + F_2(t) + F_3(t).$$

Dabei ist $F_1(t)$ eine absolut stetige Funktion, d.h.

$$F_1(t) = \int_a^t f(s) ds \quad \text{mit} \quad f(s) \geq 0.$$

Für $F_2(t)$ gibt die Summe der Sprünge $F(t_k)$, $t_k \leq t$, in den höchstens abzählbar vielen Unstetigkeitsstellen $F(t_k)$, $k = 1, 2, \dots$ an. Ein Sprung von $F(\cdot)$ bedeutet, daß die Zufallsvariable

$$\Delta Z(t_k) = \lim_{\epsilon \rightarrow 0} (Z(t_k + \epsilon) - Z(t_k - \epsilon))$$

nicht degeneriert ist, denn es gilt nach (7.4)

$$E(|Z(t_k)|^2) = \Delta F(t_k) = \lim_{\epsilon \rightarrow 0} (F(t_k + \epsilon) - F(t_k - \epsilon)) > 0.$$

Schließlich ist $F_3(t)$ eine stetige Funktion, die fast überall ¹⁾ Ableitungen gleich null besitzt. Sie liefert jedoch keinen Beitrag zum Integral (7.5) und kann deshalb in diesem Zusammenhang vernachlässigt werden. Das gleiche gilt für eine (7.9) entsprechende Zerlegung von $Z(t)$, $a \leq t \leq b$.

1) d.h.: bis auf eine Menge vom Lebesgue-Maß null.

7.2. Komplexe Form der Spektraldarstellungen

Spektraldarstellungen existieren für reelle und komplexe stationäre Prozesse mit diskretem und mit kontinuierlichem Parameterbereich. Für die Zwecke dieser Arbeit genügt es, sich auf einen diskreten Parameter zu beschränken. Wir beginnen mit der formal einfacheren Betrachtung des komplexen Falles.

Es sei $Z(\lambda)$, $-\pi \leq \lambda \leq \pi$, ein komplexwertiger Prozeß mit orthogonalen Zuwächsen. Den Parameter λ interpretieren wir aus naheliegenden Gründen als Frequenz bzw. Winkel. Wir betrachten für jeden Zeitpunkt $t = \dots, -1, 0, 1, \dots$ das stochastische Integral

$$(7.10) \quad X(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ(\lambda)$$

Dadurch ist ein diskreter i.a. komplexer Prozeß $X(t)$ definiert. Wenn $F(\lambda)$ die in (7.4) definierte Varianzfunktion von Z bezeichnet, dann erhält man als Kovarianzfunktion nach den Überlegungen im vorigen Abschnitt

$$(7.11) \quad \begin{aligned} E(X(t) \overline{X(s)}) &= \int_{-\pi}^{\pi} e^{i\lambda t} e^{-i\lambda s} E(|dZ(\lambda)|^2) \\ &= \int_{-\pi}^{\pi} e^{i\lambda(t-s)} dF(\lambda) = c(t-s) \end{aligned}$$

Wir stellen also fest, daß X ein schwach stationärer Prozeß ist. Diese Aussage ist auch umkehrbar, d.h.: Zu jedem schwach stationären Prozeß $X(t)$, $t = \dots, -1, 0, 1, \dots$ existiert ein Prozeß $Z(\lambda)$, $-\pi \leq \lambda \leq \pi$, mit orthogonalen Zuwächsen,

sodaß für X die Beziehung (7.10) und für die Kovarianzfunktion von X die Beziehung (7.11) richtig ist. Als Beweis für die Behauptung existieren verschiedene Varianten, die z.B. in (3), (7), (17), (18) oder auch in (10) zu finden sind.

Gleichung (7.10) heißt Spektraldarstellung eines schwach stationären Prozesses X . Wir wollen Z als Spektralprozeß von X bezeichnen. Gleichung (7.11) heißt Spektraldarstellung der Kovarianzfunktion eines schwach stationären Prozesses, und $F(\lambda)$ wird als spektrale Verteilungsfunktion bezeichnet¹⁾. Die Wachstumsstellen von $F(\lambda)$ nennt man das Spektrum des stationären Prozesses X . Zerlegen wir $F(\lambda)$ nach (7.9) in den absolutstetigen Anteil $F_1(\lambda)$ und den Sprunganteil $F_2(\lambda)$, dann wird aus (7.11) mit $l = t - s$:

$$(7.12) \quad c(l) = \int_{-\pi}^{\pi} e^{i\lambda l} f(\lambda) d\lambda + \sum_{k=1}^{\infty} e^{i\lambda_k l} \Delta F(\lambda_k)$$

Die entsprechende Zerlegung von (7.10) lautet

$$(7.13) \quad X(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ_1(\lambda) + \sum_{k=1}^{\infty} e^{i\lambda_k t} \Delta Z(\lambda_k)$$

Dabei ist

$$Z_1(\lambda) = Z(\lambda) - \sum_{\lambda_k < \lambda} \Delta Z(\lambda_k)$$

1) Wegen $E(|X(t)|^2) = c(0) = F(\pi) - F(-\pi) = E(|Z(\pi) - Z(-\pi)|^2)$ sagt man auch, daß durch den Prozeß Z die Varianz $c(0)$ des Prozesses X nach Maßgabe von $F(\lambda)$ in $-\pi \leq \lambda \leq \pi$ verteilt ist.

der stetige Anteil des Spektralprozesses. Bei diskretem Spektrum verschwinden die Integrale und bei stetigem Spektrum die Summen der rechten Seiten von (7.12) und (7.13).

Im letzten Fall sind die $c(l)$ die Four ierkoeffizienten der Funktion $f(\lambda)$, die zur spektralen Dichtefunktion wird. Unter gewissen Bedingungen haben wir deshalb die Fourier-entwicklung

$$(7.14) \quad f(\lambda) = \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} c(l) e^{-i\lambda l}.$$

Bei beliebigem Spektrum gilt immer eine integrierte Form von (7.14), nämlich

$$(7.15) \quad \begin{aligned} F(\lambda_2) - F(\lambda_1) &= \int_{\lambda_1}^{\lambda_2} dF(\lambda) = \\ &= \lim_{n \rightarrow \infty} \frac{1}{2\pi} \sum_{l=-n}^n c(l) \int_{\lambda_1}^{\lambda_2} e^{-i\lambda l} d\lambda \end{aligned}$$

die sich auch auf den Spektralprozeß Z übertragen läßt:

$$(7.16) \quad \begin{aligned} Z(\lambda_2) - Z(\lambda_1) &= \int_{\lambda_1}^{\lambda_2} dZ(\lambda) = \\ &= \text{l.i.m.}_{n \rightarrow \infty} \frac{1}{2\pi} \sum_{t=-n}^n X(t) \int_{\lambda_1}^{\lambda_2} e^{-i\lambda t} d\lambda \end{aligned}$$

In den oben genannten Beweisen wird gezeigt, daß ausgehend vom schwach stationären Prozeß X und dessen Kovarianzfunktion $c(\cdot)$, (7.15) und (7.16) eindeutige Definitionen für den Spektralprozeß Z und die spektrale Verteilungsfunktion $F(\cdot)$ sind.

7.5. Reelle Form der Spektraldarstellungen

Bei einem reellwertigen Prozeß X liegt es nahe, die rechte Seite von (7.10) in Realteil und Imaginärteil, der dann verschwinden muß, aufzuspalten. Mit $Z(\lambda) = U(\lambda) + iV(\lambda)$ erhält man dann:

$$\begin{aligned}
 (7.17) \quad X(t) &= \int_{-\pi}^{\pi} \cos \lambda t \, d(U(\lambda) + iV(\lambda)) + \\
 &+ i \int_{-\pi}^{\pi} \sin \lambda t \, d(U(\lambda) + iV(\lambda)) \\
 &= \int_{-\pi}^{\pi} \cos \lambda t \, dU(\lambda) - \int_{-\pi}^{\pi} \sin \lambda t \, dV(\lambda) + i(\dots) .
 \end{aligned}$$

Damit alle $X(t)$ reell sind, muß der Prozeß Z bzw. müssen die Prozesse U und V die Beziehungen

$$\begin{aligned}
 (7.18) \quad dZ(\lambda) &= \overline{dZ(-\lambda)} \quad \text{bzw.} \\
 dU(\lambda) &= dU(-\lambda) \quad \text{und} \quad dV(\lambda) = -dV(-\lambda)
 \end{aligned}$$

für $0 < \lambda \leq \pi$ erfüllen, und V darf keinen Sprung im Nullpunkt besitzen. Damit läßt sich (7.17) noch reduzieren auf:

$$X(t) = 2 \int_0^{\pi} \cos \lambda t \, dU(\lambda) - 2 \int_0^{\pi} \sin \lambda t \, dV(\lambda) - \Delta U(0) .$$

Der eventuell vorhandene Sprung $\Delta U(0) = \Delta Z(0)$ wird deshalb subtrahiert, weil er im ersten Integral der Gleichung wie die übrigen Änderungen $dU(\lambda)$ doppelt gezählt wird, in (7.17) jedoch nur einfach vorkommt. Definiert man jetzt zwei reelle Prozesse A und B durch

$$\begin{aligned}
 dA(\lambda) &= \begin{cases} dU(0) & \text{für } \lambda = 0 \\ 2 \, dU(\lambda) & \text{für } 0 < \lambda \leq \pi \end{cases} \\
 dB(\lambda) &= -dV(\lambda) \quad \text{für } 0 \leq \lambda \leq \pi ,
 \end{aligned}$$

so erhält man die Spektraldarstellung für den reellen Fall in der üblichen Fassung

$$(7.19) \quad X(t) = \int_0^{\pi} \cos \lambda t \, dA(\lambda) + \int_0^{\pi} \sin \lambda t \, dB(\lambda).$$

Aufgrund der Orthogonalität der Zuwächse $dZ(\lambda)$ und der Eigenschaften (7.18) bekommt man für die Spektralprozesse A und B:

$$(7.20) \quad \begin{aligned} E(dA(\lambda) \, dA(\mu)) &= E(dB(\lambda) \, dB(\mu)) = 0, & \lambda \neq \mu, \\ E(dA(\lambda) \, dB(\mu)) &= 0, \end{aligned}$$

$$(7.21) \quad \begin{aligned} E(dA(0))^2 &= E(|dZ(0)|^2) = dF(0), \\ E(dA(\lambda))^2 &= E(dB(\lambda))^2 = 2E(|dZ(\lambda)|^2) = 2dF(\lambda), & \lambda \neq 0. \end{aligned}$$

Die Prozesse A und B besitzen also orthogonale Zuwächse, die auch gegenseitig orthogonal sind. Ihre Varianzfunktion sei $G(\cdot)$, definiert durch

$$dG(\lambda) = \begin{cases} dF(0), & \lambda = 0 \\ 2dF(\lambda), & \lambda > 0 \end{cases}, \quad G(0) = 0.$$

$G(\cdot)$ ist die spektrale Verteilungsfunktion im reellen Fall, wie man durch Umformung von (7.11) mit den entsprechenden Überlegungen wie bei der Umformung von (7.10) feststellen kann, d.h.

$$(7.22) \quad c(t-s) = \int_{-\pi}^{\pi} \cos \lambda(t-s) \, dF(\lambda) = \int_0^{\pi} \cos \lambda(t-s) \, dG(\lambda).$$

Ohne auf weitere Einzelheiten einzugehen, fassen wir noch einmal zusammen, daß aus Symmetriegründen bei einem reellen Prozeß X zur Beschreibung der Spektraleigenschaften die Kenntnis von $Z(\lambda)$ und $F(\lambda)$ in $0 \leq \lambda \leq \pi$ genügen.

8. Lineare Funktionen stationärer Prozesse

Ist X schwach stationär, dann erfüllt auch der Prozeß

$$(8.1) \quad Y(t) = \sum_k \beta_k X(t-k), \quad \sum_k \beta_k^2 < \infty,$$

die Bedingung der schwachen Stationarität, wie man an der Spektraldarstellung

$$\begin{aligned} Y(t) &= \int_{-\pi}^{\pi} e^{i\lambda t} \sum_k \beta_k e^{-i\lambda k} dZ_X(\lambda) \\ &= \int_{-\pi}^{\pi} e^{i\lambda t} b(\lambda) dZ_X(\lambda) \end{aligned}$$

mit dem Spektralprozeß

$$(8.2) \quad dZ_Y(\lambda) = b(\lambda) dZ_X(\lambda), \quad b(\lambda) = \sum_k \beta_k e^{-i\lambda k},$$

sofort abliest. Die gleichmäßig konvergente Fourierreihe $b(\lambda)$ wird Filterfunktion genannt. Für die spektrale Verteilungsfunktion gilt

$$(8.3) \quad dF_Y(\lambda) = E(|dZ_Y(\lambda)|^2) = |b(\lambda)|^2 dF_X(\lambda),$$

sodaß sich für die Kovarianzfunktion von Y ergibt:

$$c_Y(t-s) = \int_{-\pi}^{\pi} e^{i\lambda(t-s)} |b(\lambda)|^2 dF_X(\lambda).$$

Die Funktion

$$|b(\lambda)|^2 = \sum_{k,j} \beta_k \beta_j e^{-i\lambda(k-j)}$$

heißt Übertragungsfunktion. Die Bezeichnung Filter für den

linearen Operator in (8.1), der X in Y überführt, leitet sich daraus ab, daß durch die Übertragungsfunktion das Spektrum gemäß (8.3) verstärkt bzw. abgeschwächt, also "gefiltert" wird.

Häufig kann in den Anwendungen ein Prozeß X als Summe zweier Prozesse X_1 und X_2 ausgedrückt werden, die schwach stationär und orthogonal zueinander sind, sodaß

$$(8.4) \quad E(X_1(t) \overline{X_2(s)}) = 0$$

für beliebige t und s gilt. In diesem Fall läßt sich leicht zeigen, daß die Summe $X_1 + X_2$ schwach stationär ist und die Beziehungen

$$\begin{aligned} c(t-s) &= c_1(t-s) + c_2(t-s) , \\ (8.5) \quad F(\lambda) &= F_1(\lambda) + F_2(\lambda) \quad \text{und} \\ Z(\lambda) &= Z_1(\lambda) + Z_2(\lambda) \end{aligned}$$

erfüllt sind. Schwache Stationarität liegt auch dann noch vor, wenn (8.4) von $(t-s)$ abhängt, wie man an

$$E(X(t) \overline{X(s)}) = \sum_{k,i=1}^2 E(X_i(t) \overline{X_k(s)})$$

sieht. Bei den Funktionen in (8.5) handelt es sich dann jedoch nicht mehr um die spektralen Größen von X .

9. Beispiele

Die folgenden Beispiele dienen dazu, verschiedene mathematische Formen zusammenzustellen, in denen schwach stationäre Prozesse auftreten können. Dabei sollen die Eigenschaften der zugehörigen Spektren studiert werden. Wir werden dabei immer einen reellen Prozeß X betrachten, aber je nach Bedarf sowohl die reelle als auch die komplexe Form der Spektralbeschreibungen heranziehen.

9.1. Der Prozeß aus identischen Zufallsvariablen

Darunter soll eine Folge von Zufallsvariablen $X(t)$, $t = \dots, -1, 0, 1, \dots$ verstanden werden, die alle identisch mit einer Zufallsvariablen A sind. Eine Realisation eines solchen Prozesses ist eine für alle Zeitpunkte t gleiche Konstante aus dem Wertebereich von A . Der Prozeß ist trivialerweise streng stationär und damit auch schwach stationär. Für die Kovarianzfunktion gilt

$$c(t-s) = E(X(t)X(s)) = E(A^2) = c(0) = m_A^2 + \sigma_A^2,$$

t, s beliebig.

Die spektralen Funktionen sind:

$$F(\lambda) = \begin{cases} 0, & -\pi \leq \lambda < 0 \\ c(0), & 0 \leq \lambda \leq \pi \end{cases}, \quad Z(\lambda) = \begin{cases} 0, & -\pi \leq \lambda < 0 \\ A, & 0 \leq \lambda \leq \pi \end{cases},$$

beziehungsweise

$$G(\lambda) = \begin{cases} 0, & \lambda = 0 \\ c(0), & 0 < \lambda \leq \pi \end{cases}, \quad A(\lambda) = \begin{cases} 0, & \lambda = 0 \\ A, & 0 < \lambda \leq \pi \end{cases}.$$

Dies kann durch Verifikation der Beziehungen (7.10) und (7.11) bzw. (7.19) und (7.22) leicht bestätigt werden. Bei diesem Beispiel ist bemerkenswert, daß eine Realisation des Prozesses keine Information über $E(A) = m_A$ und $\text{Var}(A) = \sigma_A^2$ liefert, sofern $\sigma_A > 0$ ist.

9.2. Prozesse mit konstantem Erwartungswert

Es sei X_0 ein schwach stationärer Prozeß mit $E(X_0(t)) = 0$ für alle t . Wir betrachten $X = X_0 + m$, sodaß $E(X(t)) = m$ für alle t gilt. Die Konstante m kann nach Beispiel 5.1. als stationärer Prozeß $A = m$ mit $\sigma_A = 0$ aufgefaßt werden. Trivialerweise ist (8.4) für X_0 und m erfüllt. Nach dem dort beschriebenen Additionsprinzip besitzt der schwach stationäre Prozeß deshalb einen Spektralprozeß

$$Z(\lambda) = \begin{cases} Z_0(\lambda), & \lambda < 0 \\ Z_0(\lambda) + m, & \lambda > 0 \end{cases}$$

der sich aus dem Spektralprozeß Z_0 von X_0 und dem in Beispiel 9.1. beschriebenen zusammensetzt. Ebenso erhält man

$$F(\lambda) = \begin{cases} F_0(\lambda), & \lambda < 0 \\ F_0(\lambda) + m^2, & \lambda > 0 \end{cases} \quad \text{sowie } c(t-s) = c_0(t-s) + m^2.$$

Die Spektralfunktionen $Z(\cdot)$ und $F(\cdot)$ von X_0 und $X = X_0 + m$ unterscheiden sich also lediglich durch einen Sprung im Nullpunkt.

9.3. Der harmonische Prozeß

Dieser Prozeß wird häufig als Standardbeispiel für die Spektraldarstellung im reellen Fall (7.19) angeführt. Man versteht darunter die Summe

$$X(t) = \sum_{k=1}^n \frac{1}{\sqrt{2}} A_k \cos \lambda_k t + \sum_{k=1}^n \frac{1}{\sqrt{2}} B_k \sin \lambda_k t$$

Da A_k und B_k sind orthogonale Zufallsvariablen mit der Eigenschaft:

$$(9.1) \quad E(A_k^2) = E(B_k^2) = \sigma_k^2.$$

Die λ_k , $k = 1, 2, \dots, n$, seien reelle Zahlen. Wegen der Periode 2π der Sinus- und Cosinusfunktion bedeutet es keine Einschränkung, wenn wir für die Frequenzen $-\pi < \lambda_k \leq \pi$ annehmen. Bei entsprechender Änderung der Vorzeichen der B_k kann man sogar davon ausgehen, daß $0 \leq \lambda_k \leq \pi$ gilt.

Die A_k und B_k sind also gerade die Sprünge der Spektralprozesse A und B in λ_k . Die σ_k^2 sind die Sprünge der spektralen Verteilungsfunktion $G(\cdot)$. Die Orthogonalität der A_k und B_k bzw. (9.1) entsprechen den Eigenschaften (7.20) bzw. (7.21).

9.4. Der stationäre Gaußsche Prozeß

Die $X(t)$ mögen normalverteilt sein mit

$$E(X(t)) = 0, \quad E(X(t)X(s)) = \begin{cases} \sigma^2 = c(0), & t = s \\ c(t-s), & t \neq s \end{cases}$$

Der Prozeß ist per def. schwach stationär. Da durch Festlegung der ersten und zweiten Momente auch die mehrdimensionale Normal-

verteilung eindeutig bestimmt ist, muß X auch streng stationär sein. Bei Gaußschen Prozessen mit $E(X(t)) = 0$ stimmen also strenge und schwache Stationarität überein. Die Variablen des Spektralprozesses haben in diesem Fall ebenfalls eine Normalverteilung mit $E(Z(\lambda)) = 0$ und die orthogonalen Zuwächse sind sogar unabhängig.

Aus dem Beispiel ist weiterhin ersichtlich, daß es zu jedem schwach stationären Prozeß X einen Gaußschen Prozeß gibt, der die gleiche Kovarianzfunktion und das gleiche Spektrum besitzt.

9.5. Der "white noise" Prozeß

Im weiteren Sinne handelt es sich hierbei um eine Folge orthogonaler Zufallsvariablen mit $E(X(t))^2 = \sigma^2$ für alle t . Im engeren Sinne sind die Variablen unabhängig und haben Erwartungswert null und die tatsächliche Varianz σ^2 . Solche Prozesse werden häufig als stochastische "Störung" in Modellen - z.B. beim linearen Regressionsmodell - eingeführt, woraus sich in Analogie zur Nachrichtentechnik die Bezeichnung "weißes Rauschen" herleitet.

Man erhält für die Kovarianzfunktion

$$c(t-s) = \begin{cases} \sigma^2, & t = s \\ 0, & t \neq s \end{cases}.$$

Damit ergibt sich die spektrale Verteilungsfunktion aus (7.15)

$$F(\lambda) = \frac{1}{2\pi} (\lambda + \pi) \sigma^2, \quad F'(\lambda) = f(\lambda) = \frac{\sigma^2}{2\pi}.$$

Das "weiße Rauschen" ist also gekennzeichnet durch ein stetiges Spektrum mit konstanter spektraler Dichtefunktion.

9.6. Der gleitende Durchschnitt

Man versteht darunter einen linear gefilterten Prozeß

$$X(t) = \sum_k \alpha_k U(t-k) \quad \text{mit} \quad \sum_k \alpha_k^2 < \infty.$$

Dabei ist U ein weißes Rauschen mit $E(U(t)) = 0$ und $\text{Var}(U(t)) = \sigma^2$.

Da U ein Spektrum $dF_U(\lambda) = \sigma^2/2\pi d\lambda$ besitzt, wie wir in Beispiel 5.5. gesehen haben, gilt nach den Überlegungen in Abschnitt 8:

$$dF_X(\lambda) = \frac{\sigma^2}{2\pi} |a(\lambda)|^2 d\lambda, \quad a(\lambda) = \sum_k \alpha_k e^{-i\lambda k}.$$

Der gleitende Durchschnitt hat also eine absolut stetige spektrale Verteilungsfunktion. In (7, S.498) wird gezeigt, daß jeder schwach stationäre Prozeß mit absolut stetigem Spektrum auch umgekehrt als gleitender Durchschnitt dargestellt werden kann. Man kann dazu jede Funktion $a(\lambda)$ wählen, die

$$|a(\lambda)|^2 = \frac{2\pi}{\sigma^2} f_X(\lambda)$$

erfüllt. Die Entwicklung dieser Funktion in eine Fourierreihe

$$a(\lambda) = \sum_{k=-\infty}^{\infty} \alpha_k e^{-i\lambda k}$$

liefert die Koeffizienten des gleitenden Durchschnitts. Die α_k können reell gewählt werden, wenn X reell ist. Die Kovarianzfunktion des gleitenden Durchschnitts kann explizit durch

$$\begin{aligned} E(X(t)X(s)) &= \sum_{k,j} \alpha_k \alpha_j E(U(t-k)U(t-j)) = \\ &= \sum_{k,j} \alpha_k \alpha_j \delta_{t-k, t-j} = c(t-s) \end{aligned}$$

bestimmt werden.

9.7. Der autoregressive Prozeß

Ein stochastischer Prozeß heißt autoregressiver Prozeß q -ter Ordnung, wenn er der stochastischen Differenzengleichung

$$(9.2) \quad \beta_0 X(t) - \beta_1 X(t-1) - \dots - \beta_q X(t-q) = U(t)$$

genügt, und die Bedingung

$$(9.3) \quad E(U(t)X(s)) = 0, \quad \text{für } s < t$$

erfüllt. U ist wie im Beispiel 9.5. ein orthogonaler Prozeß mit $E(U(t))^2 = \sigma^2$. Ohne Beschränkung der Allgemeinheit können wir $\beta_0 = 1$ wählen. Wenn wir uns einen "Anfangszustand" $X(0), \dots, X(-q+1)$ unabhängig von $U(t)$, $t = 1, 2, \dots$ vorgeben, dann ist klar, daß

$$X(t) = \beta_1 X(t-1) + \dots + \beta_q X(t-q) + U(t)$$

für $t = 1, 2, \dots$ ein autoregressiver Prozeß ist. Die Größe der β_k spielt dabei keine Rolle. Wir fragen nun, wann der autoregressive Prozeß stationär ist. Man kann zeigen, daß dies genau dann der Fall ist, wenn die charakteristische Gleichung

$$z^q - \beta_1 z^{q-1} - \dots - \beta_q = 0$$

nur Wurzeln mit $|z_1| < 1$ hat. Das ist gleichbedeutend mit der Aussage, daß die homogene Lösung der Differenzengleichung

gedämpft ist. Ist diese Bedingung erfüllt, dann können wir die in Abschnitt 6 beschriebene Filtermethode anwenden. Für die bekannte Spektralverteilung von U muß gelten

$$dF_u(\lambda) = \frac{\sigma^2}{2\pi} d\lambda = |b(\lambda)|^2 dF_x(\lambda)$$

Der stationäre autoregressive Prozeß hat also eine absolutstetige Spektralverteilung mit der Dichte

$$f(\lambda) = \frac{\sigma^2}{2\pi} |b(\lambda)|^{-2}$$

Dabei lautet die Filterfunktion

$$b(\lambda) = 1 - \sum_{k=1}^a B_k e^{-i\lambda k}$$

Zurückgreifend auf das Beispiel 9.6 erhalten wir als zusätzliches Ergebnis, daß jeder autoregressive stationäre Prozeß auch als gleitender Durchschnitt

$$X(s) = \sum_k \alpha_k U(s-k), \quad \sum_k \alpha_k^2 < \infty$$

dargestellt werden kann. Wenden wir darauf (9.3) an, dann ergibt sich weiter für $k > 0$:

$$E(U(s+k)X(s)) = \alpha_{-k} \sigma^2 = 0, \quad \text{d.h.} \quad \alpha_{-k} = 0.$$

Schließlich führt die Kombination von (9.2), (9.3) und $B_0 = 1$ auch noch auf

$$\sigma^2 = E(X(t)U(t)) = \alpha_0 \sigma^2 \quad \text{d.h.} \quad \alpha_0 = 1.$$

Der autoregressive Prozeß ist demnach gleich einem einseitigen

gleitenden Durchschnitt

$$(9.4) \quad X(t) = \sum_{k=0}^{\infty} \alpha_k U(t-k), \quad \alpha_0 = 1,$$

Die Berechnung der α_k kann elementar durchgeführt werden. Wir gehen aus von

$$(9.5) \quad X(t) = \sum_{k \geq 1} \beta_k X(t-k) + U(t)$$

und setzen darin $X(t-1)$ rekursiv ein. Wir erhalten

$$\begin{aligned} X(t) &= \beta_1 \sum_{k \geq 1} \beta_k X(t-1-k) + \sum_{k \geq 2} \beta_k X(t-k) + \\ &\quad + U(t) + \beta_1 U(t-1) = \\ &= \sum_{k \geq 2} (\beta_k + \beta_1 \beta_{k-1}) X(t-k) + U(t) + \beta_1 U(t-1). \end{aligned}$$

Setzt man dieses Verfahren schrittweise fort, dann bekommt man für jedes $m \geq 1$ eine Darstellung

$$(9.6) \quad X(t) = \sum_{k \geq m+1} \beta_{k,m+1} X(t-k) + \sum_{j=0}^m \beta_{j,j} U(t-j),$$

$$\beta_{0,0} = 1,$$

wobei die $\beta_{k,m}$ rekursiv durch

$$\beta_{k,1} = \beta_k, \quad k \geq 1 \quad (\text{mit } \beta_k = 0, \quad k > q)$$

$$\beta_{k,m+1} = \beta_{k,m} + \beta_{m,m} \beta_{k-1,m}, \quad k \geq m+1, \quad m = 1, 2, \dots$$

definiert sind. Die Diagonalelemente $\beta_{m,m}$ in dieser Entwicklung ergeben die α_m des gleitenden Durchschnitts (9.4).

9.8. Ein Beispiel von Doob

Doob (7, S. 479) betrachtet das folgende, verblüffend einfache Beispiel eines stationären Prozesses. Es sei Λ eine Zufallsvariable, die eine Verteilung in $-\pi \leq \Lambda \leq \pi$ mit der Verteilungsfunktion $F(\lambda)$ besitzt und r eine reelle Zahl. Wir definieren durch

$$(9.7) \quad X(t) = r \cdot e^{i\Lambda t} \quad t = \dots, -1, 0, 1, \dots$$

einen (komplexen) stochastischen Prozeß. Dieser ist stationär, denn

$$E(X(t)\overline{X(s)}) = r^2 \int_{-\pi}^{\pi} e^{i\lambda(t-s)} dF(\lambda) = c(t-s)$$

Man sieht, daß $F(\lambda)$ bis auf die Konstante r^2 mit der spektralen Verteilungsfunktion des Prozesses X übereinstimmt. Der Zufallsmechanismus, der eine Realisation dieses Prozesses erzeugt, besteht einfach darin, eine bestimmte Frequenz λ zu wählen. Die Amplitude ist für alle Realisationen gleich. Im Beispiel 9.3 haben wir den harmonischen Prozeß kennengelernt, der einen ganz anderen Mechanismus zur Erzeugung von Realisation besitzt. Zum Vergleich mit (9.7) wollen wir die komplexe Form

$$(9.8) \quad X(t) = \sum_{j=1}^n R_j e^{i\lambda_j t} \quad \text{mit } E(R_j \overline{R_k}) = 0, j \neq k$$

betrachten. Wenn wir davon ausgehen, daß die R_j auch unabhängig sind, dann besteht eine Realisation von (9.8) in den unabhängigen Realisationen von gewissen r_j .

Nun ist es offensichtlich, daß man durch die Wahl einer geeigneten diskreten Verteilung von Λ auf den λ_k von (9.8) und durch geeignete Wahl von r^2 erreichen kann, daß die Spektralen Verteilungsfunktionen von (9.8) und (9.7) übereinstimmen. Nach

den obigen Überlegungen sind die Realisationen aber ganz verschiedener Arten von trigonometrischen Reihen. "Thus the spectral distribution function does not determine the harmonic analysis of the individual sample..." (7 , S. 479). Diese Aussage gilt natürlich auch für reelle Prozesse X . Die reelle Variante von (9.8) wird dabei beschrieben durch

$$(9.9) \quad X(t) = r \cos(\Lambda t + \phi)$$

mit Λ beliebig verteilt in $0 \leq \Lambda \leq \pi$ und ϕ gleichmäßig verteilt in $0 \leq \phi \leq 2\pi$, Λ und ϕ unabhängig.

5.9. Ein Beispiel von Nerlove

Zur Untersuchung der Saisonkomponente benutzt Nerlove (35 , S. 261) ein Modell, das mit (9.9) verwandt ist. Es lautet

$$X(t) = R(t) \cos(\lambda t + \phi) \quad t = \dots, -1, 0, 1, \dots$$

Die Zufallsvariable ϕ ist wie in (9.9) definiert und R ist selbst ein schwach stationärer Prozeß mit Erwartungswert null. Die $R(t)$ sind unabhängig von ϕ . Als Kovarianzfunktion erhält man wegen der Unabhängigkeit

$$\begin{aligned} E(X(t)X(s)) &= E(R(t)R(s)) E(\cos(\lambda t + \phi) \cos(\lambda s + \phi)) \\ &= c_R(t-s) \frac{1}{2\pi} \int_0^{2\pi} \cos(\lambda t + \phi) \cos(\lambda s + \phi) d\phi \\ &= c_R(t-s) \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2} (\cos \lambda(t-s) + \cos(\lambda(t+s)+2\phi)) d\phi \end{aligned}$$

$$= c_R(t-s) \frac{1}{2} \cos \lambda(t-s) = c_X(t-s) .$$

Der Prozeß ist also schwach stationär. Da die Amplitude zu jedem Zeitpunkt eine andere sein kann, handelt es sich bei den Realisationen jedoch nicht um reine Sinus- oder Cosinusschwingungen. Es kommt auf die Art der Stationarität an, inwieweit sich in den Realisationen Schwingungen mit einer festen Frequenz erkennen lassen.

III. Lineare Prognose bei stationären Prozessen

10. Das Prognoseproblem

Die allgemeine Aufgabenstellung für die Prognose eines (reellen) stochastischen Prozesses besteht darin, auf der Basis von gewissen "Vergangenheitswerten" $X(t-m), \dots, X(t-n)$, $0 < m < n$, eine Prognose für den Zeitpunkt t zu erstellen. Dabei soll die Prognose, die wir mit $X(t, m, n)$ bezeichnen wollen, zunächst eine beliebige Funktion $g(\cdot)$ der Vergangenheitswerte sein. Beschränken wir uns auf das Prinzip der Minimierung der Varianz, dann muß also

$$E(X(t) - g(X(t-m), \dots, X(t-n)))^2$$

bezüglich aller Funktionen $g(\cdot)$ minimiert werden. Die Lösung dieser Aufgabe ist der bedingte Erwartungswert

$$(10.1) \quad E(X(t) \mid X(t-m), \dots, X(t-n))$$

der bekanntlich minimale Varianz besitzt. Ohne Kenntnis der gemeinsamen Verteilungsfunktion der auftretenden Variablen - diese Situation liegt bei schwach stationären Prozessen vor, bei denen wir ja nur die zweiten Momente der Verteilung kennen - kann (10.1) jedoch nicht berechnet werden. Wir betrachten deshalb das einfachere lineare Problem, den Ausdruck

$$E((X(t) - \sum_{k=m}^n b_k X(t-k))^2$$

bezüglich der b_k zu minimieren.

Differentiation und Nullsetzen ergibt die Normalgleichungen

$$- 2E(X(t-k) (X(t) - \sum_{j=m}^n b_j X(t-j))) = 0, k = m, m+1, \dots, n$$

Ist X stationär, dann wird daraus

$$- 2c(k) + 2 \sum_{j=m}^n b_j c(k-j) = 0, k = m, m+1, \dots, n$$

oder vereinfacht und ausführlich geschrieben

$$b_m c(0) + b_{m+1} c(-1) + \dots + b_n c(m-n) = c(m)$$

$$(10.2) \quad b_m c(1) + b_{m+1} c(0) + \dots + b_n c(m-n+1) = c(m+1)$$

$$b_m c(n-m) + b_{m+1} c(n-m-1) + \dots + b_n c(0) = c(n)$$

Bei reellem Prozeß X ist natürlich $c(k-j) = c(j-k)$.

Als beste lineare Prognose erhält man

$$(10.3) \quad X(t, m, n) = \sum_{k=m}^n b_k X(t-k)$$

wobei die $b_k = b_k(m, n)$ die Lösungen des Gleichungssystems (10.2) sind und nicht vom Zeitpunkt t abhängen. Ebenso ist die Varianz des Prognosefehlers

$$\sigma^2(m, n) = E(X(t) - X(t, m, n))^2$$

nur von der Kovarianzfunktion $c(\cdot)$ abhängig, aber nicht vom Zeitpunkt t.

Die Varianz des Prognosefehlers kann offenbar nicht zunehmen, wenn wir die Prognosebasis erweitern. Deshalb gelten die Ungleichungen

$$(10.3) \quad \sigma(m,n) \geq \sigma(m,n+1), \quad \sigma(m+1,n) \geq \sigma(m,n) .$$

Aus der ersten Ungleichung folgt, daß mit $n \rightarrow \infty$ die Varianzen monoton fallend gegen einen Wert $\sigma(m) \geq 0$ konvergieren. Über die Theorie der Hilberträume kann man zeigen¹⁾, daß auch die $X(t,m,n)$ im quadratischen Mittel gegen eine Zufallsvariable

$$X(t,m) = \lim_{n \rightarrow \infty} X(t,m,n)$$

konvergieren, die als beste lineare Prognose auf der Basis der Vergangenheitswerte $X(t-k)$, $k \geq m$ aufgefaßt werden kann. Es gelten nach wie vor die Normalgleichungen

$$(10.4) \quad E(X(t-k) (X(t) - X(t,m))) = 0, \quad k \geq m .$$

Die Prognose $X(t,m)$ kann i.a. aber nicht mehr als Summe der $X(t-k)$, $k \geq m$ wie in (10.3) dargestellt werden. Dies gilt jedoch z.B., wenn der Prozeß autoregressiv ist. Die Definitionsgleichungen (9.5) bzw. (9.6) bilden bei diesem Prozeßtyp gleichzeitig die Zerlegung in die Prognose

$$X(t,m) = \sum_{k \geq m} \beta_{k,m} X(t-k) = E(X(t) | X(t-k), k \geq m)$$

und den Prognoserest

$$X(t) - X(t,m) = \sum_{j=0}^{m-1} \beta_{j,j} U(t-j) .$$

1) Vgl. z.B. Urbanik (47)

Die Normalgleichungen (10.4) folgen per def. aus (9.3).
Gleichzeitig stellen wir fest, daß die beste lineare Prognose
gleich dem bedingten Erwartungswert ist, sodaß die beiden am
Anfang besprochenen Prognoseaufgaben zusammenfallen.

11. Reguläre und deterministische Prozesse

Man betrachte für $m = 1, 2, \dots$ die verschiedenen Prognosen
 $X(t, m)$ und die zugehörigen Varianzen $\sigma(m)$ der Prognosefehler.
Aus der zweiten Ungleichung von (10.3) ergibt sich mit $n \rightarrow \infty$
die Ungleichungskette

$$\sigma(1) \leq \sigma(2) \leq \sigma(3) \leq \dots$$

Es können nun zwei Fälle eintreten

a) $\sigma(1) \neq 0$, d.h. es kann kein einziger zukünftiger Wert genau
vorausgesagt werden. Da dies der Normalfall in den Anwendungen
sein dürfte, wird der Prozeß regulär genannt.

b) $\sigma(1) = 0$, d.h. $X(t)$ ist genau prognostizierbar durch $X(t-1)$,
 $X(t-2)$ Durch rekursive Anwendung dieser von t unabhängigen
Aussage ergibt sich, daß $X(t)$ auch durch $X(t-m)$, $X(t-m-1)$...
genau vorausgesagt werden kann. Aus $\sigma(1) = 0$ folgt also $\sigma(m)$
 $= 0$ für alle m . Ein Prozeß, der $\sigma(1) = 0$ erfüllt, wird deshalb
deterministisch genannt.

Ein triviales Beispiel eines deterministischen Prozesses ist
eine Folge von identischen Zufallsvariablen wie in 9.1 be-
schrieben. Der schon wiederholt betrachtete autoregressive

Prozeß ist ein Beispiel eines regulären Prozesses. Aus der Schilderung im vorigen Abschnitt ergibt sich, daß $U(t)$ aus der Definition (9.5) gerade der Prognosefehler der Einschrittprognose ist mit der Varianz

$$\sigma^2(1) = \text{Var} (U(t)) .$$

Reguläre und deterministische Prozesse können auch durch ihre spektrale Verteilungsfunktion charakterisiert werden. Wir kommen darum noch einmal auf das Prognoseproblem zurück, indem wir den Prognosefehler

$$X(t) - X(t,1,n) = X(t) - \sum_{k=1}^n b_k X(t-k)$$

als linear gefilterten Prozeß mit der Filterfunktion

$$b(\lambda) = 1 - \sum_{k=1}^n b_k e^{-i\lambda k}$$

auffassen. Wir erhalten mit den Gleichungen des 8. Abschnitts

$$(11.1) \quad X(t) - X(t,1,n) = \int_{-\pi}^{\pi} e^{i\lambda t} \left(1 - \sum_{k=1}^n b_k e^{-i\lambda k} \right) dZ(\lambda)$$

und

$$(11.2) \quad E(X(t) - X(t,1,n))^2 = \int_{-\pi}^{\pi} \left| 1 - \sum_{k=1}^n b_k e^{-i\lambda k} \right|^2 dF(\lambda)$$

Mit W bezeichnen wir das Spektrum von $F(\lambda)$, d.h. die Menge derjenigen Punkte in $-\pi \leq \lambda \leq \pi$, in denen $F(\lambda)$ wächst und mit S das Komplement von W .

Beachten wir die Zerlegung von $F(\lambda)$ entsprechend (7.9) in den absolut stetigen Anteil

$$F_1(\lambda) = \int_{-\pi}^{\lambda} f(u) du$$

und den Sprunganteil $F_2(\lambda)$, dann sieht man, daß W aus den Sprungstellen von $F_2(\lambda)$ bzw. $F(\lambda)$ und denjenigen Punkten besteht, in denen $f(\lambda) > 0$ ist, während S die Punkte enthält, für die $f(\lambda) = 0$ ist und gleichzeitig $\Delta F(\lambda) = \Delta F_2(\lambda) = 0$ ist.

Die Frage nach der besten Prognose ist nun gleichbedeutend mit dem Problem, den Betrag der Filterfunktion auf W möglichst klein zu machen. Das ist besonders einfach, wenn W nur aus endlich vielen Sprungstellen von $F(\lambda)$ besteht, das Spektrum also diskret ist. Wir bezeichnen die Sprungstellen mit λ_j , $j=1,2,\dots,n$. Wenn X reell ist - und davon sind wir bisher beim Prognoseproblem ausgegangen - dann liegen die Sprungstellen wegen der Symmetrie von $\Delta F(\lambda)$ symmetrisch zum Nullpunkt. Wir betrachten nun ein komplexes Polynom

$$P(z) = \prod_{j=1}^n (e^{-i\lambda_j} - z) = \sum_{k=0}^n b'_k z^k.$$

Da die Wurzeln $e^{-i\lambda_j}$ dieses Polynoms konjugiert komplex auftreten oder gleich 1 sind - letzteres wenn ein $\lambda_j = 0$ ist - werden die Koeffizienten b'_k alle reell. Insbesondere erhält man für b'_0 durch Zusammenfassung der konjugiert komplexen Wurzeln

$$b'_0 = \prod_{j=1}^n e^{-i\lambda_j} = e^{i\sum \lambda_j} = 1.$$

Indem wir $b_0 = b'_0$ und $b_k = -b'_k$, $k=1,2,\dots,n$ setzen, bilden wir eine Filterfunktion

$$b(\lambda) = 1 - \sum_{k=1}^n b_k e^{-i\lambda k} = P(e^{-i\lambda}),$$

die für das Prognoseproblem geeignet ist. Der Wert des Integrals (11.2) muß gleich null werden, da $b(\lambda)$ auf W - den Sprungstellen - verschwindet und $F(\lambda)$ nach Voraussetzung auf S konstant ist. Wir haben damit explizit gezeigt, daß Prozesse mit diskretem und endlichem Spektrum deterministisch sind. Die beste Prognose hängt nur von endlich vielen Vergangenheitswerten $X(t-k)$, $k=1,2,\dots,n$ ab und kann mittels der b'_k angegeben werden.

Bei beliebigem Spektrum funktioniert die geschilderte Konstruktionsmethode zur Berechnung der b_k nicht. Man sieht aber aus (11.2), daß die Varianz $\sigma^2(1)$ des Fehlers der Prognose $X(t,1)$, die alle Vergangenheitswerte berücksichtigt, gleich dem Infimum des Integrals sein muß, genommen über alle endlichen Folgen $b_1, b_2, \dots, b_n$, $n=1,2,\dots$, d.h.:

$$\sigma^2(1) = \inf_{b_1, \dots, b_n} \int_{-\pi}^{+\pi} \left| 1 - \sum_{k=1}^n b_k e^{-i\lambda k} \right|^2 dF(\lambda) .$$

Der Prozeß X ist also genau dann deterministisch, wenn das Infimum gleich null ist. In der Prognosetheorie¹⁾ wird gezeigt, daß dies wiederum genau dann der Fall ist, wenn

$$(11.3) \quad \int_{-\pi}^{+\pi} \log f(\lambda) d\lambda = -\infty$$

gilt, wobei $f(\lambda)$ die Dichte des absolut stetigen Teils der spektralen Verteilungsfunktion ist. Daraus ergibt sich, daß keinesfalls nur die stationären Prozesse mit diskretem Spektrum, für die (11.3) trivialerweise mit $f(\lambda)$ identisch null erfüllt wird, deterministisch sind. Das Spektrum kann sogar absolut stetig sein. Mit Hilfe der Komplementärmenge S des Spektrums W kann man als hinreichendes Kriterium formu-

¹⁾ Vgl. Urbanik (47)

lieren: Der Prozeß X ist deterministisch, wenn S ein ganzes Intervall umfaßt, d.h. wenn $F(\lambda)$ auf einem wenn auch noch so kleinen Intervall konstant ist. Umgekehrt kann bei einem regulären Prozeß die Menge S kein solches Intervall enthalten.

Die Bedeutung der Begriffe "deterministisch" und "regulär" sollte in der Praxis nicht überbewertet werden. Sie sind ja an das lineare Prognoseproblem geknüpft. Im Hinblick auf das allgemeine Prognoseproblem könnte man einen Prozeß auch deterministisch nennen, wenn

$$X(t) = E(X(t)|X(t-1), X(t-2), \dots)$$

gilt. Das Beispiel 9.8 ist in diesem Sinne deterministisch, denn man erhält

$$E(X(t)|X(t-1), X(t-2), \dots) = E(re^{i\lambda t} | \Lambda) = re^{i\lambda t}.$$

Im Sinne der Definition b) ist das Beispiel jedoch nur dann deterministisch, wenn (11.3) erfüllt ist. Man beachte dabei, daß $F(\lambda)$ in 9.8 eine vollkommen beliebige Verteilungsfunktion sein kann.

12. Die Zerlegung von Wold

Die regulären stationären stochastischen Prozesse sind der Gegenstand der sogenannten Wold-Zerlegung. Jeder solcher Prozeß $X(t)$, $t = \dots, -1, 0, 1, \dots$ kann nach Wold (53) zerlegt werden in

$$(12.1) \quad X(t) = \sum_{k=0}^{\infty} \alpha_k U(t-k) + V(t), \quad \sum |\alpha_k|^2 < \infty, \quad \alpha_0 = 1.$$

Dabei ist U ein weißes Rauschen, orthogonal zu V und es gilt

$$(12.2) \quad E(X(s)U(t)) = 0, \quad s < t, \quad E(U(t))^2 = \sigma^2(1)$$

mit $\sigma(1)$ gleich der Varianz des Prognosefehlers für $m=1$. Die Zerlegung kann mit Hilfe der in Abschnitt 6 beschriebenen linearen Prognose ($m=1$) explizit angegeben werden.

$$U(t) = X(t) - X(t,1) \quad \text{mit} \quad E(U(t))^2 = \sigma^2(1)$$

$$(12.3) \quad \alpha_k = E(X(t) U(t-k)) / \sigma^2(1), \quad k \geq 0$$

$$V(t) = X(t) - \sum_{k=0}^{\infty} \alpha_k U(t-k) \quad .$$

Die Orthogonalitätseigenschaft (12.2) ergibt sich als eine Folgerung aus den Normalgleichungen (10.4) für n gegen unendlich; die Orthogonalität der $U(t)$ selbst ebenfalls. Daß U und V orthogonal sind, kann mit (12.3) direkt verifiziert werden, falls man die Konvergenz der rechten Seite in der Gleichung für $V(t)$ einmal voraussetzt, also annimmt, daß die Summe der Quadrate der α_k beschränkt ist. Letzteres ist ein Ergebnis, das nur durch einen differenzierteren Beweis, bei dem u.a. Hilbertraumtheorie benutzt wird, erbracht werden kann ¹⁾. Demnach ist V ein schwach stationärer deterministischer Prozeß. Seine Spektralschar besteht aus den Sprüngen der Spektralschar von X . Der Prozeß V entspricht unter der Voraussetzung der Regularität also dem Summenanteil in der Zerlegung (7.13). Der einseitige gleitende Durchschnitt

$$Y(t) = \sum_{k=0}^{\infty} \alpha_k U(t-k)$$

ist der Integralanteil jener Zerlegung. Sein Spektrum ist absolut stetig, wie wir im Beispiel 9.6 gesehen haben. Die spektrale Ver-

1) Vgl. Doob (7) und Urbanik (47)

teilungsfunktion ist der absolut stetige Anteil der spektralen Verteilungsfunktion von X . Der Prozeß Y ist regulär. Die Varianzen der Prognosefehler von Y sind die gleichen wie die von X , da ja V deterministisch ist. Nun ist die m -Schritt-Prognose für $Y(t)$ wegen der Orthogonalität der $U(t)$ durch

$$Y(t, m) = \sum_{k=m}^{\infty} \alpha_k U(t-k)$$

gegeben, so daß man als Varianz der Prognosefehler erhält

$$\begin{aligned} \sigma_x^2(m) &= \sigma_y^2(m) = E(Y(t) - Y(t, m))^2 = \\ &= E\left(\sum_{k=0}^{m-1} \alpha_k U(t-k)\right)^2 = \left(\sum_{k=0}^{m-1} \alpha_k^2\right) \sigma^2(1). \end{aligned}$$

Die regulären Prozesse mit absolut stetigem Spektrum werden auch vollkommen undeterministisch genannt. Wie man an der Wold-Zerlegung sieht, können sie als einseitige gleitende Durchschnitte dargestellt werden. Da andererseits jeder einseitig gleitende Durchschnitt auch regulär ist¹⁾, stellt dieser Prozeßtyp den einzigen vollkommen undeterministischen dar. Er ist durch ein absolut stetiges Spektrum mit

$$\int_{-\pi}^{+\pi} \log f(\lambda) d\lambda > -\infty$$

gekennzeichnet.

Bemerkenswert an der Wold-Zerlegung ist noch, daß der deterministische Anteil ein diskretes Spektrum besitzt, was nicht für alle deterministischen Prozesse zutrifft, wie im Abschnitt 11 festzustellen war.

1) Vgl. Doob (7, S. 569)

IV. Statistische Grundlagen der Spektralanalyse

Das Anliegen der spektralanalytischen Untersuchung einer Zeitreihe besteht darin, aufgrund der Spektraleigenschaften Aussagen über die Art des stochastischen Prozesses zu liefern, der diese Zeitreihe erzeugt. Setzt man die Kenntnis der spektralen Funktionen voraus, dann werden diese Aussagen durch das Wahrscheinlichkeitstheoretische Tautologiengebäude geliefert, dessen möglichst vollständige Beschreibung in den ersten beiden Kapiteln versucht wurde.

Wir wollen uns nun mit dem Problem befassen, wie denn aus einer beobachteten Zeitreihe $x_1, x_2, \dots, x_n$ auf die spektralen Eigenschaften geschlossen werden kann. Wir gehen dabei davon aus, daß sich aus a-priori-Kenntnissen entscheiden läßt, ob die Zeitreihe Realisation eines schwach stationären Prozesses ist oder nicht. Diese Frage kann nur so und nicht etwa durch einen statistischen Test geklärt werden, da bekanntermaßen jede endliche Zahlenreihe als trigonometrisches Polynom darstellbar ist und demnach als Realisation eines entsprechenden stationären harmonischen Prozesses wie in Beispiel 5.3. angesehen werden kann.

13. Ergodische Eigenschaften stationärer Prozesse

Für den schwach stationären Prozeß X , den wir auch weiterhin als reell annehmen, mögen Realisationen im Zeitintervall $t=1, 2, \dots, n$ vorliegen. Wir wollen zunächst die Stichprobenfunktion

$$(13.1) \quad \frac{1}{n} \sum_{t=1}^n X(t) e^{-i\lambda t} = R_n(\lambda)$$

betrachten.

Sie wird nahegelegt, wenn man sich die Aufgabe stellt, die n Zeitreihenwerte durch ein trigonometrisches Polynom auszu-
drücken. Man hat dann bekanntermaßen an den Stellen $\lambda_j = +2\pi j/(n-1)$,
 $j=0,1,\dots,(n-1)/2$ - dabei sei n o.B.d.A. ungerade - die Amplituden
 $P_j = P_n(\lambda_j)$ zu berechnen. Man kann beweisen¹⁾, daß die Stichproben-
funktion (13.1) den Grenzwert

$$(13.2) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n X(t) e^{-i\lambda t} = \Delta Z(\lambda)$$

besitzt mit der Eigenschaft

$$(13.3) \quad E(|\Delta Z(\lambda)|^2) = \Delta F(\lambda) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=0}^{n-1} c(l) e^{-i\lambda l}.$$

Dabei sind wie gehabt $Z(\lambda)$ und $F(\lambda)$ der Spektralprozeß bzw. die
spektrale Verteilungsfunktion. Der Grenzwert (13.2) ist fast über-
all gleich null und nur in den abzählbar vielen Sprungstellen des
Spektrums gleich einer von null verschiedenen Zufallsvariablen. Über
diese Zufallsvariable können wir mit einer einzigen Stichprobe aller-
dings keine Aussage erhalten, es sei denn, der Sprung des Spektral-
prozesses ist entartet und besitzt keine Varianz. Ist dies der Fall
und gilt also die Konvergenz (13.2) für fast alle Realisationen, dann
wollen wir den Prozeß individuell ergodisch in λ nennen. Die ent-
sprechende diskrete Komponente hat dann keinen stochastischen Charakter
mehr. In der besonderen Situation, wenn der stochastische Prozeß X
konstante Erwartungswerte hat und individuell ergodisch ist für
 $\lambda = 0$, wird X mittelwertergodisch genannt.

Eine weitere Stichprobenfunktion, die sich zur Untersuchung eines
stationären Prozesses sofort anbietet, ist die erwartungstreue
Schätzung

1) Vgl. Doob (7, S.489)

$$(13.4) \quad \tilde{c}(1) = \frac{1}{n-1} \sum_{l=1}^{n-1} X(t) X(t+1), \quad l=0,2,\dots,n-1$$

für die Kovarianzfunktion. Wir sagen, daß der Prozeß kovarianzergodisch sei, wenn (13.4) mit Wahrscheinlichkeit 1 gegen $c(1)$ konvergiert. Um Kriterien für die Konvergenz zu erhalten, wird gewöhnlich angenommen, daß mit $X(t)$ auch das Produkt $X(t)X(t+1)$ für jeden festen Abstand l ein schwach stationärer Prozeß ist, was bei strikt stationären Prozessen automatisch gilt. Es stellt sich heraus¹⁾, daß X unter diesen Umständen genau dann kovarianzergodisch ist, wenn die vierten Momente die Bedingung

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n E(X(0) X(1) X(t) X(t+1)) = c(1)$$

erfüllen. Das kann so interpretiert werden, daß die in der Klammer stehenden Variablen $X(0) \cdot X(1)$ und $X(t) \cdot X(t+1)$, die ja beide den Erwartungswert $c(1)$ haben, im Durchschnitt nicht miteinander korreliert sind oder für größer werdende t eine immer kleiner werdende Korrelation aufweisen. Dies wird um so eher plausibel sein, wenn auch die Variablen $X(0)$ und $X(t)$ für große t unkorreliert sind. Man kann zeigen, daß bei einem Gaußschen Prozeß mit $E(X(t)) = 0$ die Ergodizität äquivalent ist mit

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{l=0}^n c^2(l) = 0.$$

Letzteres ist wiederum genau dann erfüllt, wenn die spektrale Verteilungsfunktion stetig ist. An dem folgenden Gegenbeispiel ist zu sehen, daß diese Aussage nicht auf schwach stationäre Prozesse schlechthin übertragen werden kann.

1) Vgl. Doob (7, S. 493)

Es sei X ein harmonischer Prozeß (vgl. Beispiel 9.3) definiert durch

$$X(t) = A \cos \phi t + B \sin \phi t, \quad E(A) = E(B) = E(AB) = 0,$$

$$P(A=a) = P(B=a) = P(A=-a) = P(B=-a) = \frac{1}{2},$$

wobei a eine reelle Zahl größer null ist. Als Produkt zweier Variablen zu verschiedenen Zeitpunkten berechnen wir

$$\begin{aligned} X(t) X(t-1) &= A^2 \cos \phi t \cos \phi(t-1) + B^2 \sin \phi t \sin \phi(t-1) \\ &\quad + AB (\cos \phi t \sin \phi(t-1) + \sin \phi t \cos \phi(t-1)). \end{aligned}$$

Nun ist wegen der Konstruktion $A^2 = B^2 = a^2$. Deshalb und aufgrund bekannter Additionstheoreme bekommt man

$$\frac{1}{n-1} \sum_{t=1}^{n-1} X(t) X(t-1) = a^2 \cos \phi + \frac{1}{n-1} \sum_{t=1}^{n-1} AB \sin \phi(2t-1).$$

Der Kosinusausdruck auf der rechten Seite ist die zu schätzende Kovarianz $c(1)$, während die Sinusreihe für n gegen unendlich verschwindet. Der Prozeß ist also kovarianzergodisch, obwohl seine spektrale Verteilungsfunktion einen Sprung in $\lambda = \phi$ macht.

14. Das Periodogramm

14.1. Wiener's Periodogrammbegriff

Die Namen Periodogramm und integriertes Periodogramm sind von Schuster in (43) und von Wiener in (51) eingeführt worden. Wegen

der unterschiedlichen Bedeutung dieser Größen, je nachdem welcher Zeitreihenbegriff zu Grunde gelegt wird, ist es notwendig, kurz auf die Darstellung Wieners und die Art der Zeitreihen einzugehen, die er zum Ausgangspunkt seiner Betrachtung macht¹⁾.

Zunächst muß darauf hingewiesen werden, daß Wiener sich nicht auf stochastische Prozesse bezieht, sondern einen Zufallscharakter außer Acht läßt und die Zeitreihe einfach als eine Funktion der (als Kontinuum gedachten) Zeit t definiert. In unserer Terminologie wäre dies also eine Realisation eines stochastischen Prozesses. Der Wahrscheinlichkeitstheoretische Hintergrund spielt jedoch wie gesagt überhaupt keine Rolle. Die wesentliche Voraussetzung - formuliert für eine diskrete Zeitreihe $X(t)$, $t = \dots, -1, 0, 1, \dots$ - lautet vielmehr, daß für jede ganze Zahl l

$$(14.1) \quad \frac{1}{2n} \sum_{t=-n}^n X(t) X(t+l)$$

mit $n \rightarrow \infty$ einen Grenzwert besitzt. Bezeichnet man diesen Grenzwert mit $c(l)$, dann wird sowohl die Funktion

$$(14.2) \quad \lim_{n \rightarrow \infty} \frac{1}{2n} \sum_{l=-n}^n c(l) e^{-i\lambda t}$$

von λ als auch deren stetige Approximation

$$(14.3) \quad \frac{1}{4n^2} \left| \sum_{t=-n}^n X(t) e^{-i\lambda t} \right|^2 = r_n^2(\lambda)$$

als Periodogramm bezeichnet. Schließlich definiert Wiener auch

1) Vgl. dazu (51, S. 126ff.)

ein "integriertes Periodogramm" durch

$$(14.4) \quad \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} c(l) \frac{e^{-i\lambda l} - 1}{il} = \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} c(l) \int_0^1 e^{-i\phi l} d\phi.$$

Die Größen (14.2) - (14.4) sind im Sinne der stationären Prozesse interpretierbar, wenn der Prozeß kovarianzergodisch ist. Die $c(l)$ entsprechen natürlich den Werten der Kovarianzfunktion. Das Periodogramm im Wienerschen Sinne (14.2) ist im wesentlichen mit der Gleichung (13.3) identisch und also eine Funktion, die nur in den Sprungstellen des Spektrums von Null verschieden ist, aber keine Aussagen über den stetigen Teil liefert. Das ist auch der Grund, weswegen Wiener das integrierte Periodogramm eingeführt hat, das bis auf eine Konstante mit der spektralen Verteilungsfunktion übereinstimmt, wie man an (7.15) sehen kann. Der Periodogrammbegriff (14.3) schließlich entstammt dem klassischen Problem der diskreten harmonischen Analyse. Die Bedeutung spezieller Werte für gewisse λ_j wurde bei (13.1) schon beschrieben.

14.2. Der Periodogrammbegriff bei stochastischen Prozessen

In der neueren Literatur, die sich mit der Schätzung der spektralen Dichtefunktion

$$f(\lambda) = \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} c(l) e^{-i\lambda l}$$

eines schwach stationären Prozesses X befaßt, dessen spektrale Verteilungsfunktion absolut stetig ist, wird etwas abweichend von (14.3)

die Stichprobenfunktion

$$(14.5) \quad P_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^n X(t) e^{-i\lambda t} \right|^2$$

als Periodogramm bezeichnet. Das läßt sich umformen in

$$(14.6) \quad P_n(\lambda) = \frac{1}{2\pi} \sum_{l=-n+1}^{n-1} \bar{c}(l) e^{-i\lambda l}$$

mit

$$(14.7) \quad \bar{c}(l) = \frac{1}{n} \sum_{t=1}^{n-l} X(t) X(t+l) ,$$

woraus man sofort ablesen kann, daß $P_n(\lambda)$ den asymptotischen Erwartungswert $f(\lambda)$ hat. Für den Fall eines Gaußschen Prozesses kann man außerdem als asymptotische Varianz berechnen

$$\lim_{n \rightarrow \infty} \text{Var}(P_n(\lambda)) = \begin{cases} f^2(\lambda) & \text{für } \lambda \neq 0 \\ 2f(0) & \text{für } \lambda = 0 \end{cases} .$$

Das Periodogramm ist also keinesfalls eine konsistente Schätzfunktion für $f(\lambda)$. Eine andere Möglichkeit, die zu einer konsistenten Schätzung führt, ist das integrierte Periodogramm, wie wir es im Zusammenhang mit stationären Prozessen durch

$$(14.8) \quad \int_{\lambda_1}^{\lambda_2} P_n(\lambda) d\lambda = \frac{1}{2\pi n} \int_{\lambda_1}^{\lambda_2} \left| \sum_{t=1}^n X(t) e^{-i\lambda t} \right|^2 d\lambda$$

definieren wollen.

Doob (7, S. 496) zeigt, daß allein unter der Voraussetzung, daß der Prozeß kovarianzergodisch ist, in allen Punkten λ_1, λ_2 , in denen $F(\lambda)$ stetig ist, für fast alle Realisationen die Beziehung

$$\lim_{n \rightarrow \infty} \int_{\lambda_1}^{\lambda_2} P_n(\lambda) d\lambda = F(\lambda_2) - F(\lambda_1)$$

erfüllt ist. Das gilt auch dann, wenn $F(\lambda)$ Sprungstellen zwischen λ_1 und λ_2 hat. Wir können also feststellen, daß bei stetigem Spektrum die spektrale Dichtefunktion nicht konsistent durch das Periodogramm geschätzt werden kann, daß aber die Zuwächse der spektralen Verteilungsfunktion konsistent durch das integrierte Periodogramm geschätzt werden können, auch dann, wenn die spektrale Verteilungsfunktion unstetig ist.

15: Konsistente Schätzung der spektralen Dichte

15.1. Das gewichtete Periodogramm

Die Überlegungen der beiden vorhergehenden Abschnitte machen einsichtig, daß man sich bei der Schätzung des Spektrums sinnvollerweise auf stationäre Prozesse beschränkt, die kovarianzergodisch sind, denn ohne diese Voraussetzung können die zentralen Stichprobengrößen (13.4) oder (14.7) nicht als geschätzte Werte der Kovarianzfunktion interpretiert werden. Üblicherweise wird auch angenommen, daß die spektrale Verteilungsfunktion absolut stetig sei, was bei Gaußschen Prozessen durch die erste Annahme schon impliziert ist. Es gibt mehrere Möglichkeiten, unter diesen Voraussetzungen konsistente Schätzfunktionen zu konstruieren. Sehr ausführlich ist dieses Problem von Grenander und Rosenblatt (17) untersucht worden. Wir wollen uns hier auf die Betrachtung einer Klasse von Schätzfunktionen

$$(15.1) \quad \overline{f}(\lambda) = \int_{-\pi}^{\pi} P_n(\lambda - \phi) W_n(\phi) d\phi$$

beschränken, die ein gewichtetes Periodogramm darstellen und fast ausschließlich in den Anwendungen auftauchen. Das Periodogramm $P_n(\cdot)$ muß man sich dabei über $-\pi$ und π hinaus periodisch fortgesetzt denken. Die Funktion $W_n(\phi)$ ist eine Gewichtsfunktion, die sich um den Nullpunkt konzentriert und die Eigenschaften

$$(15.2) \quad W_n(-\phi) = W_n(\phi), \quad \int_{-\pi}^{\pi} W_n(\phi) d\phi = 1$$

besitzt. Sie heißt Spektralfenster. Ein Beispiel ist das sogenannte Rechteckfenster

$$(15.3) \quad W_n(\phi) = \begin{cases} \frac{1}{2h} & , -h < \phi < h, \quad h = h(n), \\ 0 & \text{sonst} \end{cases}$$

Damit die Schätzung (15.1) asymptotisch erwartungstreu ist, muß gelten

$$(15.4) \quad f(\lambda) = \lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} f(\lambda - \phi) W_n(\phi) d\phi.$$

Die Gleichung erhält man durch Grenzübergang aus (15.1) unter der Beachtung der asymptotischen Erwartungstreue von $P_n(\lambda)$. Wenn (15.4) für beliebige spektrale Dichten richtig sein soll, muß $W_n(\phi)$ gegen die Dirac-funktion konvergieren; d.h. außer (15.2) ist

$$(15.5) \quad \lim_{n \rightarrow \infty} \int_{\phi > \varepsilon} |W_n(\phi)| d\phi = 0 \quad \text{für alle } \varepsilon > 0$$

eine notwendige und hinreichende Bedingung für asymptotische Erwartungstreue.

Die Bedingungen für die Konsistenz einer Schätzung (15.1) sind schwieriger herzuleiten. Wir beschränken uns hier darauf festzustellen, daß das Spektralfenster für $n \rightarrow \infty$ "nicht zu schnell" gegen die Dichtefunktion konvergieren darf, weil sonst die Schätzung fast identisch ist mit dem Periodogramm, das wir als nicht-konsistent erkannt haben. Um diese Bedingung zu erfüllen, sind eine Reihe von Vorschlägen für geeignete Spektralfenster gemacht worden. Sie beziehen sich auf Fenster, welche in eine Fourierreihe

$$(15.6) \quad W_n(\phi) = \frac{1}{2\pi} \sum_{k=-n}^n w_n(k) e^{-i\phi k} = \frac{w_n(0)}{2\pi} + \frac{1}{\pi} \sum_{k=1}^n w_n(k) \cos \phi k$$

entwickelt werden können. Setzt man (15.6) und die Entwicklung (14.6) von $P_n(\lambda)$ in (15.1) ein, multipliziert aus und integriert gliedweise, dann erhält man

$$(15.7) \quad \bar{f}_n(\lambda) = \frac{1}{2\pi} \sum_{k=-n}^n w_n(k) \bar{c}(k) e^{-i\lambda k} = \frac{\bar{c}(0) w_n(0)}{2\pi} + \frac{1}{\pi} \sum_{k=1}^n w_n(k) \bar{c}(k) \cos \lambda k.$$

Die Folge der Fourierkoeffizienten $w_n(k)$, $k=1,2,\dots$ werden in diesem Zusammenhang als lag-Fenster bezeichnet. Man sieht, daß es auf die Werte für $k=0$ überhaupt nicht ankommt. Die Bedingung (15.5) impliziert aber für $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} w_n(k) = 1 \quad \text{für alle } k.$$

Der Forderung der nicht allzu schnellen Konvergenz des Spektralfensters $W_n(\phi)$ gegen die Dichtefunktion wird nun dadurch Rechnung getragen, daß man in (15.7) nur die ersten Glieder berücksichtigt und eine gegen n relativ kleine ganze Zahl m wählt, oberhalb der die $w_n(k) = 0$ sind. Die

Abhängigkeit des Spektralfensters von n wird außerdem reduziert auf die Abhängigkeit von dieser ganzen Zahl m . Das Spektralfenster lautet dann $W_m(\phi)$ und das lag-Fenster $w_m(k)$. Die oben erwähnten Vorschläge unterscheiden sich in den $w_m(k)$ für $k > m$. Ihnen ist gemeinsam, daß sie zu konsistenten Schätzfunktionen für die spektrale Dichte führen, wenn m mit n gegen unendlich geht, gleichzeitig aber das Verhältnis m/n gegen null strebt.

15.2. Einige Spektralfenster

Wir wollen im Überblick einige Fenster und ihre Eigenschaften in Anlehnung an (17, S. 145ff.) zusammenstellen. Neben der Konsistenz ist dabei für die praktische Anwendung vor allem der mittlere quadratische Fehler

$$E(\bar{f}_n(\lambda) - f(\lambda))^2 = B_n^2(\lambda) + \text{Var}(\bar{f}_n(\lambda))$$

wichtig, der sich in das Quadrat der Verzerrung

$$B_n(\lambda) = E(\bar{f}_n(\lambda)) - f(\lambda)$$

und die Varianz von $\bar{f}_n(\lambda)$ aufspalten läßt.

Das einfachste lag-Fenster, das die genannte Bedingung erfüllt ist

$$w_m(k) = 1 \quad \text{für } k \leq m \quad .$$

Die zugehörige Schätzfunktion ist ein "abgeschnittenes" Periodogramm, das sich nur durch die kleinere Anzahl der Summanden von (14.6) unterscheidet. Das Spektralfenster hat die Form

$$(15.8) \quad W_n(\phi) = \frac{1}{2\pi} \frac{\sin(m + 1/2)\phi}{\sin \phi/2} \quad .$$

Verzerrung und Varianz haben für "große" m und n die Werte¹⁾

$$(15.9) \quad B_n(\lambda) \approx \frac{1}{2\pi m} \int_{-\pi}^{\pi} \frac{\sin \phi}{1 - \cos \phi} f'(\lambda - \phi) d\phi$$

$$\text{Var}(\bar{f}(\lambda)) \approx a \frac{m}{n} f^2(\lambda)$$

mit $a = 2$, wenn $\lambda \neq 0$ und $a = 4$, wenn $\lambda = 0$ ist.

Auf Bartlett geht das lag-Fenster

$$(15.10) \quad w_m(k) = \left(1 - \frac{k}{m}\right) \quad \text{für } k \leq m$$

zurück, das zu einer Schätzfunktion mit einer Verzerrung führt, die im Vergleich zu (15.9) um den Faktor n/m vergrößert ist und deren Varianz um das 2/3-fache verkleinert ist. Das Spektralfenster für (15.10) lautet

$$W_m(\phi) = \frac{1}{2\pi} \frac{\sin^2 m\phi/2}{\sin^2 \phi/2}.$$

Es hat im Vergleich zum Spektralfenster (15.8) keine "negativen Seitenfenster".

Von Tukey stammt der Vorschlag

$$w_m(k) = 0.46 \cos k\pi/m + 0.54 \quad \text{für } k \leq m,$$

der für "große" m und n auf die Verzerrung und die Varianz

1) Hier und im folgenden nehmen wir natürlich an, daß die auftretenden Ableitungen der spektralen Dichte auch existieren.

$$B_n(\lambda) \approx 0,23 (\pi/m)^2 f''(\lambda)$$

$$\text{Var}(\tilde{f}(\lambda)) \approx a \frac{m}{2n} f^2(\lambda) (0,46^2 + 2 \cdot 0,54^2)$$

führt. Der Faktor a hat die gleiche Bedeutung wie in (15.9). Man sieht, daß hier die Varianz im Vergleich zu (15.9) rund um den Faktor 0.4 verkleinert ist, allerdings auf Kosten der Regularitätsanforderungen an die spektrale Dichte.

Auch das Rechteckfenster (15.3) führt auf eine konsistente Schätzung der Dichte, wenn die halbe Fensterbreite h in Abhängigkeit von n geeignet gewählt wird. Setzen wir $h = \pi/m$, wobei m wie oben eine Zahl sein soll, die mit n gegen unendlich geht, aber das Verhältnis m/n gegen null streben läßt, dann nimmt das lag-Fenster die Werte

$$\begin{aligned} w_m(k) &= \int_{-\pi}^{+\pi} w_m(\phi) e^{i\phi k} d\phi \\ &= \frac{1}{2h} \int_{-h}^h \cos \phi k d\phi = 1 \quad \text{für } k = 0 \\ &= \frac{1}{hk} \sin hk = \frac{m}{\pi k} \sin k\pi/m \quad \text{für } k \neq 0 \end{aligned}$$

an. Für die Verzerrung der Schätzung ergibt sich

$$B_n(\lambda) \approx (\pi/m)^2 f''(\lambda) / 6 .$$

Die Varianz ist die Hälfte derjenigen von (15.9).

An den vorgestellten Beispielen wird die Problematik der Schätzung der spektralen Dichtefunktion deutlich. Es ergibt sich ein Zielkonflikt, wenn man versucht, gleichzeitig die Varianz und die Verzerrung der Schätzung zu verkleinern. Bei vorgegebenem Stichprobenumfang n ist das eine nurauf Kosten des anderen möglich.

15.3. Prewhitening und Recoloring

Die Verfahrensweise, die sich hinter diesem sehr suggestiven Namen mehr oder weniger verbirgt, kann folgendermaßen beschrieben werden: Es sei X der stationäre Prozeß mit absolut stetigem Spektrum, dessen spektrale Dichte $f_x(\lambda)$ geschätzt werden soll. Ausgehend von der Tatsache, daß insbesondere die Verzerrung $B_n(\lambda)$ sehr stark von den Regularitätseigenschaften in der Nähe des Schätzpunktes abhängt, kann man versuchen, den Prozeß X in einen Prozeß Y überzuführen, der eine möglichst konstante Dichtefunktion $f_y(\lambda)$ hat (Prewhitening). Geschieht diese Transformation durch einen linearen Filter

$$Y(t) = \sum_k B_k X(t-k)$$

dann besteht zwischen den beiden Dichtefunktionen nach (8.3) die Beziehung

$$(15.11) \quad f_y(\lambda) = |b(\lambda)|^2 f_x(\lambda)$$

mit der Filterfunktion $b(\lambda)$. Nach der Schätzung der Dichte von Y kann über (15.11) dann eine Schätzung für die Dichte von X angegeben werden (Recoloring).

Die methodischen Schwierigkeiten des Verfahrens liegen auf der Hand. Nur wenn man a-priori-Kenntnisse über den Verlauf der Dichtefunktion hat und die Filterfunktion entsprechend wählen kann, ist gewährleistet, daß der transformierte Prozeß eine zumindest in bestimmten Bereichen glatt verlaufende Dichtefunktion besitzt. Das Verfahren ist deshalb zur Schätzung einer vollkommen unbekannten Dichte nicht geeignet.

V. Spektralanalyse ökonomischer Reihen

16. Motivationen für den Ansatz

Im Kapitel IV. wurde schon dargestellt, daß spektralanalytische Untersuchungen im Prinzip aus zwei Teilen bestehen. Erstens aus der Bestimmung der Spektraleigenschaften aufgrund einer Stichprobe und zweitens aus den sich daraus ergebenden Folgerungen für den Prozeß aufgrund der stochastischen Theorie stationärer Prozesse.

Bei statistischen Anwendungen tritt der zweite Teil mehr oder weniger in den Hintergrund. Als Ziel der Spektralanalyse wird z.B. bei O. Morgenstern (33) genannt "... to discover whether periodic fluctuations occur, where they lie, and to measure their energy ". In den Anwendungen sind die Grundlage dafür die Spektraldarstellungen eines schwach stationären Prozesses, der ergodisch ist und eine spektrale Dichtefunktion besitzt. Es dreht sich also um die Beziehungen

$$(16.1) \quad X(t) = \int_{-\pi}^{+\pi} e^{i\lambda t} dZ(\lambda), \quad |dZ(\lambda)|^2 = f(\lambda).$$

$$c(l) = \int_{-\pi}^{\pi} e^{i\lambda l} f(\lambda) d\lambda$$

$$f(\lambda) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} c(l) e^{-i\lambda l} dl.$$

Bei der so verstandenen Spektralanalyse werden andere Modelle schwach stationärer Prozesse, wie wir sie in den Kapiteln II. und III. kennengelernt haben, nicht in Betracht gezogen.

Es stellt sich die Frage, ob die in (16.1) enthaltenen Voraussetzungen erfüllt sind. Zu den Fürsprechern, die diese Frage bejahen und sich

ausdrücklich für die Anwendung der Spektralanalyse bei ökonomischen Reihen einsetzen, gehört wiederum O. Morgenstern (33). Er beschreibt die Situation einer Volkswirtschaft folgendermaßen: "Producers and consumers, being subject to very regular, in many cases strictly periodic, influences, send signals of these activities into the (linear) time invariant "black box" called "economy". Their intermingling and the impinging of chance influences produces the aggregative data which confronts the economist."

Um von dieser verbalen Beschreibung auf stationäre Prozesse als Erklärung ökonomischer Zeitreihen zu kommen, bedarf es allerdings noch einiger Ergänzungen. Die "Aktivitäten", von denen hier die Rede ist, müssen quantifizierbar sein, und die "Regelmäßigkeiten" müssen in Sinus- und Cosinusschwingungen ausgedrückt werden, so daß eine einzelne Aktivität, die mit einer bestimmten Periode T oder Frequenz $\lambda = 2\pi/T$ auftritt, als Input

$$X_{\lambda}(t) = A \cos \lambda t + B \sin \lambda t$$

des "Schwarzen Kastens" geschrieben werden kann. Da die Volkswirtschaft linear wirken soll, ist der Output von dergleichen Form und die Summe aller Aktivitäten - gekennzeichnet durch beliebige Frequenzen und zugehörige Amplituden - ergibt einen Gesamtoutput

$$(16.2) \quad X(t) = \sum_j A_j \cos \lambda_j t + B_j \sin \lambda_j t$$

Wenn man dabei auf kontinuierlich viele Aktivitäten abstellen will, kann man (16.2) auch durch

$$(16.3) \quad X(t) = \int_{-\infty}^{\infty} \cos \lambda t dA(\lambda) + \int_{-\infty}^{\infty} \sin \lambda t dB(\lambda)$$

ersetzen, wobei die Zuwächse $dA(\lambda)$ und $dB(\lambda)$ den A_j und B_j entsprechen. Damit ist man jedoch noch nicht am Ziel eines stationären stochastischen

Prozesses angelangt. Dazu müssen die Aktivitäten als zufällig und unabhängig interpretiert werden. Berücksichtigt man außerdem, daß der Prozeß nur in diskreten Zeitpunkten $t = \dots, -1, 0, 1, \dots$ beobachtet wird, dann können alle Aktivitäten mit einer Periode $T < 1$ bzw. mit einer Frequenz $\lambda > 2\pi$ auch als Aktivitäten angesehen werden, deren Periode ein ganzzahliges Vielfaches von T ist, das gerade größer als 1 ist. Diese Problematik wird in der Literatur mit "Aliasing" bezeichnet. Formal kann also für einen diskreten Prozeß Aktivitäten mit Perioden > 1 unterstellt werden. Benutzt man jetzt noch Symmetrieeigenschaften, wie im Abschnitt 9 am Beispiel des harmonischen Prozesses dargestellt, dann bekommt man aus (16.3) nach entsprechender Änderung der Prozesse A und B die reelle Spektraldarstellung

$$(16.4) \quad X(t) = \int_0^{\pi} \cos \lambda t \, dA(\lambda) + \int_0^{\pi} \sin \lambda t \, dB(\lambda)$$

als Modell für eine ökonomische Zeitreihe. Daß diese Darstellung äquivalent ist mit einer komplexen Form wie in (16.1), haben wir schon wiederholt festgestellt.

Es bedarf keiner großen Erläuterungen, daß die Motivation dieses Ansatzes ökonomisch nur wenig fundiert ist. Wenn man den Ausgangspunkt akzeptiert, daß eine Volkswirtschaft linear auf Inputs einwirkt, und daß die Inputs zum großen Teil sich gleichförmig und regelmäßig wiederholen, lassen sich auch andere gleichbedeutende Zeitreihenmodelle aufstellen. Beispielsweise genügt die Summe harmonischer Schwingungen (16.2) unter sehr geringen Voraussetzungen an die nichtstochastischen Größen A_j und B_j der Bedingung, daß (14.1) konvergiert und ist damit der deskriptiv gleichwertigen Theorie Wieners zugänglich. Formel (16.1) ist also keinesfalls zwingend und stellt nur eine Möglichkeit zur Untersuchung zyklischen Verhaltens dar.

17. Die Stationarität der Komponenten

17.1. Allgemeines

Gegenüber Zeitreihenmodellen, die auf den vier Komponenten Trend, Zyklus, Saison und irregulärer Rest aufbauen, hat der spektrale Ansatz den Vorteil, daß gleichzeitig alle möglichen Frequenzen berücksichtigt werden. Wie sind die vier Komponenten aber in dieses Modell einzubeziehen? Es ist plausibel, die Jahresfrequenz - z.B. $\lambda = 2\pi/12$ bei Monatsdaten oder $\lambda = 2\pi/4$ bei Vierteljahresdaten - und deren Vielfaches der Saisonkomponente zuzuordnen. Alle Frequenzen, die kleiner als die Jahresfrequenz sind - die Perioden sind dann größer als ein Jahr - werden z.B. nach Nullau (36) den zyklischen Schwankungen zugerechnet. Eine Ausnahme bilden die Schwingungsbeiträge mit sehr kleinen Frequenzen nahe null, die den Trend ausmachen sollen. Trend und zyklische Komponente, d.h. alle Frequenzen zwischen null und der Jahresfrequenz werden zur glatten Komponente zusammengefaßt. Alle übrigen Frequenzen schließlich gehören zur irregulären Komponente.

Die oben vorgeschlagene Integration der vier Komponenten einer Zeitreihe im Rahmen des Spektralansatzes (16.1) bzw. (16.4) ist nicht in jeder Hinsicht befriedigend. Besonders der als Schwingung aufgefaßte Trend steht im Widerspruch zur herkömmlichen Vorstellung. Wir müssen deshalb den oben mehr oder weniger plausibel gemachten Ansatz genauer untersuchen.

Zu diesem Zweck zerlegen wir den Frequenzbereich $-\pi \leq \lambda \leq \pi$ des Ansatzes (16.1) in disjunkte Teilbereiche I_1, I_2, I_3 und I_4 , die, wie oben beschrieben, den vier Komponenten zukommen.¹⁾ Entsprechend ergibt sich eine Zerlegung der Spektralschar

1) Man beachte, daß im komplexen Ansatz die Frequenzbereiche I_k symmetrisch zum Nullpunkt liegen.

$$Z(\lambda) = Z_1(\lambda) + Z_2(\lambda) + Z_3(\lambda) + Z_4(\lambda)$$

$$Z_k(\lambda) = \begin{cases} Z(\lambda) & \text{für } \lambda \in I_k \\ 0 & \text{sonst} \end{cases},$$

woraus die Aufspaltung des Prozesses X in

$$X(t) = X_1(t) + X_2(t) + X_3(t) + X_4(t)$$

$$X_k(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ_k(\lambda) = \int_{I_k} e^{i\lambda t} dZ(\lambda)$$

resultiert. Die vier Prozesse $X_k(t)$, die den vier Komponenten entsprechen, sind selbst schwach stationär und orthogonal zueinander. Außerdem wissen wir aus Abschnitt 8, wo die Summe orthogonaler Prozesse betrachtet wurde, daß sich die Kovarianzfunktionen und die Spektralverteilungsfunktionen ebenfalls addieren. Wenn also die ökonomische Reihe Realisation eines schwach stationären Prozesses ist, dann gilt dies auch für die vier Komponenten. Wir sind deshalb in der Lage, die Stationaritätseigenschaften von $X(t)$ in einfacher Weise auf die Stationaritätseigenschaften der einzelnen Komponenten zurückzuführen.

17.2. Der Trend

In Abschnitt 6 hatten wir bereits gesehen, daß die Erwartungswerte $E(X(t))$ des Prozesses durch die Kovarianzfunktion beschränkt werden. Die üblichen Vorstellungen, die den Trend z.B. als linear oder exponentiell wachsende Bewegung des Erwartungswertes eines stochastischen Prozesses sehen, sind demnach im Spektralansatz nicht integrierbar. Nur in einem begrenzten Zeitabschnitt kann eine streng monotone Entwicklung durch die Realisation eines schwach stationären

nären Prozesses beschrieben werden. Beispielsweise zeigt eine Realisation des "Trends"

$$X_1(t) = A \sin \lambda t \quad \text{mit } E(A)=0 \text{ und } \text{Var}(A)=\sigma^2$$

eine fast lineare Entwicklung für kleine Werte von t auf, die umso länger abhält, je kleiner die Frequenz λ ist, und mit umso größerer Wahrscheinlichkeit einen steilen Verlauf annimmt, je größer die Varianz von A ist. Diese Eigenschaft, daß der Trendverlauf im spektralanalytischen Ansatz durch den Zufall und die Varianz der Spektralschar erklärt werden muß, ist typisch und steht im Gegensatz zu den bekannten Trendbeschreibungen. Außerdem stellen wir fest, daß diese Beschreibung für eine unbeschränkt wachsende Entwicklung langfristig nicht anwendbar ist. Deswegen ist auch die Frage unerheblich, ob der "Spektraltrend" X_1 ein stetiges Spektrum besitzt und ergodisch ist, denn diese Eigenschaften kommen erst langfristig zum tragen.

17.3. Die Saisonkomponente

Zu den Standardeigenschaften der Saisonkomponente gehört, daß ihre Werte periodisch wiederkehren und daß sie als Abweichung vom Durchschnitt über eine Grundperiode anzusehen sind. Üblicherweise wählt man als entsprechendes Modell eine Fourierreihe

$$(17.1) \quad X_2(t) = \sum_{k=1}^{T/2} (A_k \cos \lambda_k t + B_k \sin \lambda_k t) \quad \text{mit } \lambda_k = \frac{2\pi}{T} k,$$

wobei T o.B.d.A. als gerade Zahl angenommen wird. Bei Interpretation der Amplituden A_k und B_k als orthogonale Zufallsvariablen, wie im Beispiel des harmonischen Prozesses, erhalten wir Stationarität.

Allerdings stellt (17.1) noch einen Prozeß mit diskretem Spektrum dar. Es stellt sich die Frage, ob man auch ein stetiges Spektrum vor- aussetzen kann. Von der rein deskriptiven Seite, d.h. wenn man nur die endlich vielen Werte einer Zeitreihe im Auge hat, läßt sich diese Frage bejahen. Man kann sich die Punktmassen in den diskreten Fre- quenzstellen auch in einer beliebig kleinen Umgebung dieser Stellen verteilt denken. Von der Interpretationsseite her muß man die Frage aber verneinen. Um das einzusehen, nehmen wir der Einfachheit halber an, daß die spektrale Saisonkomponente X_3 konstante Erwartungswerte gleich null besitzt. Ist $x_3(t)$ ein bestimmter realisierter Wert, dann wird auch $x_3(t+jT)$ mit T als Grundperiode der Saison sehr in der Nähe liegen, gleichgültig wie groß j ist. Das bedeutet aber, daß die Variablen zu diesen Zeitpunkten hoch korreliert sind, und daß man sinnvollerweise für die Kovarianzen

$$c(T) = c(jT) \quad \text{für } j = 1, 2, \dots$$

annehmen sollte. Zumindest kann man nicht davon ausgehen, daß $c(jT)$ gegen Null geht für j gegen unendlich. Unter diesen Umständen kann die Saisonkomponente aber keine spektrale Dichte besitzen, denn für eine solche Funktion gilt

$$\lim_{t \rightarrow \infty} \int_{-\pi}^{\pi} e^{i\lambda t} f(\lambda) d\lambda = 0$$

Dieses Integral stellt aber nach (7.11) gerade den Wert $c(t)$ der Kovarianzfunktion dar. Die spektrale Verteilungsfunktion der Saisonkomponente $X_2(t)$ besitzt also Sprungstellen. Wir bemerken aber, daß dennoch die spektrale Saisonkomponente kovarianzergodisch sein kann, wie wir an dem in Abschnitt 13 beschriebenen Gegenbeispiel gesehen haben.

17.4 Die zyklische Komponente

Bei der zyklischen Komponente einer ökonomischen Reihe kann man davon ausgehen, daß der entsprechende spektrale Teilprozeß $X_3(t)$ einen Erwartungswert konstant gleich null besitzt. In dieser Annahme kommt nämlich zum Ausdruck, daß man i.a. die zyklische Komponente als Abweichung vom Trendverlauf auffaßt, und daß man ohne eine Kenntnis der "konjunkturellen" Situation zu einem Zeitpunkt t auch keine Aussage zugunsten irgendeiner Abweichung treffen kann. Natürlich ist der bedingte Erwartungswert $E(X_3(t) | X_3(t-1), X_3(t-2) \dots)$ von null verschieden. Auch kann man mit einiger Überzeugungskraft voraussetzen, daß die Korrelation zwischen zwei Variablen zu verschiedenen Zeitpunkten nur vom Abstand der Zeitpunkte voneinander abhängen wird¹⁾, und daß die Korrelation mit zunehmenden Abstand gegen Null gehen wird. Der Unterschied zur Saisonkomponente, der hierin zum Ausdruck kommt, liegt darin begründet, daß man bei der Saisonkomponente eine exakte Periode hat, nach der im Prinzip ähnliche Verhältnisse auftreten, während bei der zyklischen Komponente eine solche exakte Periode nicht auftritt. Damit ist die Annahme der Stationarität gerechtfertigt, und wenn die Konvergenz der Kovarianz $c(d)$ gegen null genügend schnell ist, wird die spektrale Verteilungsfunktion auch absolut stetig sein. Da zumindest bei Gaußschen Prozessen die Stetigkeit der spektralen Verteilungsfunktion notwendig und hinreichend dafür ist, daß der Prozeß auch kovarianzergodisch ist, spricht schließlich auch nichts gegen diese letzte Eigenschaft. Die notwendige Voraussetzung zur vernünftigen Schätzung einer spektralen Dichtefunktion der zyklischen Komponente einer ökonomischen Reihe wäre damit erfüllt.

1) Auch hier muß man natürlich wieder von der Kenntnis der konjunkturellen Situation - z.B. Aufschwung oder Abschwung - absehen.

Daß sich die irreguläre Komponente $X_{11}(t)$ als schwach stationärer und ergodischer Prozeß mit Erwartungswert gleich null und absolut stetigem Spektrum angesehen werden kann, bedarf keiner näheren Rechtfertigung.

18. Die Ausschaltung von Komponenten

18.1. Filtermethoden

Aus dem Für und Wider bezüglich der schwachen Stationarität der einzelnen Komponenten hat sich ergeben, daß die zyklische- und die irreguläre Komponente uneingeschränkt die Voraussetzung zur Schätzung der spektralen Dichtefunktion erfüllen. Für die beiden anderen Komponenten trifft dies nicht zu. Je nachdem ob man sich auf die Beschreibung einer endlichen Zeitreihe beschränkt oder langfristige asymptotische Aspekte berücksichtigt, kommt man zu dem einen oder anderen Ergebnis. Aus dieser Situation ergibt sich unter entsprechenden Umständen die Notwendigkeit, vor der Schätzung der spektralen Dichtefunktion eine Trend- oder Saisonbereinigung durchzuführen. Die Bereinigung der Zeitreihe von Saison- oder Trendeinflüssen kann auch der eigentliche Zweck einer Zeitreihenanalyse sein.

Eine der Möglichkeiten, nämlich die Anwendung von linearen Filtern, haben wir in Abschnitt 8 schon kennengelernt. Zur Erklärung der Wirkung von Filtermethoden muß der Prozeß schwach stationär sein. Er braucht aber weder ergodisch zu sein, noch muß er eine spektrale Dichtefunktion besitzen. Als Beispiel zur Saisonausschaltung bei Vierteljahresdaten betrachten wir den gleitenden Durchschnitt

$$\begin{aligned}
 Y(t) &= \frac{1}{8} (X(t-2) + 2 \sum_{k=-1}^1 X(t+k) + X(t+2)) \\
 &= \sum_{k=-2}^2 \beta_k X(t+k)
 \end{aligned}$$

mit der Filterfunktion

$$\begin{aligned}
 b(\lambda) &= \sum_{k=-2}^2 \beta_k e^{i \cdot k \lambda} \\
 &= \frac{1}{4} (1 + 2 \cos \lambda + \cos 2\lambda) ,
 \end{aligned}$$

die reellwertig ist, weil der gleitende Durchschnitt symmetrisch ist. In $0 \leq \lambda \leq \pi$ hat $b(\lambda)$ die Nullstellen $\pi/2$ und π . Das entspricht der Jahresperiode 4 und der Halbjahresperiode 2, die bei Vierteljahresdaten die Saisonkomponente ausmachen. Da nach (8.3) die Beziehung

$$dF_Y(\lambda) = |b(\lambda)|^2 dF_X(\lambda)$$

gilt, werden also die Sprungstellen $\Delta F_X(\pi/2)$ und $\Delta F_X(\pi)$ und damit die Saisonkomponente ausgeschaltet. Aber auch das Wachstum im übrigen Teil des Spektrums wird verändert, denn nur für $\lambda = 0$ ist $|b(\lambda)|^2 = 1$, so daß es sich bei $Y(t)$ keinesfalls um einen nur saisonbereinigten Prozeß handelt.

Zur Trendbereinigung werden unter anderem sogenannte Hochpaßfilter verwendet, die so geartet sein müssen, daß die Übertragungsfunktion möglichst gut den Eigenschaften

$$|b(\lambda)|^2 = \begin{cases} 0 & \text{für } \lambda \text{ nahe } 0 \\ 1 & \text{sonst} \end{cases}$$

entspricht.

Die Subtraktion gleitender Durchschnitte relativ großer Länge von den Originaldaten liefert solche Filter.¹⁾

18.2. Interpretationsprobleme

Beim Filtern einer Zeitreihe zum Zwecke der Ausschaltung gewisser Komponenten tritt der Nebeneffekt auf, daß das Spektrum auch in unerwünschten Bereichen verändert wird. Ein berühmtes Beispiel dafür sind die europäischen Weizenpreise der Jahre 1500 - 1869, die Lord Beveridge 1922 einer Periodogrammanalyse unterworfen hat.²⁾ Da die Reihe, die zwischen den Werten 17 im Jahre 1500 und 224 im Jahre 1869 schwankt, einen ausgesprochenen Trend aufweist, der die Amplitude der zyklischen und irregulären Bewegung beeinflußt, berechnete Beveridge aus den Preisen $p(t)$ einen "Fluktuationsindex"

$$X(t) = P(t) / \frac{1}{31} \sum P(t+k) = P(t) / \bar{P}(t)$$

der stationäres Verhalten aufweist.³⁾ Übereinstimmend mit Beveridge stellen Granger und Hughes fest, daß das geschätzte Spektrum von $X(t)$ ein relatives Maximum für eine Periode von 15 Jahren hat. Andererseits zeigt aber das geschätzte Spektrum von $P(t)$ ein relatives Maximum für eine Periode von 13,3 Jahren. Die von Granger und Hughes untersuchte Frage, ob nun die Zeitreihe eine zyklische Komponente von 15 Jahren oder 13,3 Jahren enthält, darf allerdings nicht überbewertet werden, denn bei den beiden geschätzten Maxima ist die jeweils andere Frequenz nur wenig unterhalb des Gipfels anzutreffen. Dennoch

1) Vgl. König und Wolters (25, S.100)

2) Granger und Hughes (16)

3) Die numerischen Werte der Indices findet man in Kendall und Stuart (22, S. 406), eine graphische Darstellung bei Granger und Hughes (16).

gehen Granger und Hughes davon aus, daß die Periode von 13.3 die in den Ausgangsdaten eigentlich vorhandene sei, während die von 15.2 Jahren als Folge eines systematischen Effekts auftritt, wie er von Hochpaßfiltern bekannt ist¹⁾ und den sie in erster Näherung auch für das Trendbereinigungsverfahren von Beveridge nachweisen. Methodisch ist dieser Nachweis allerdings nicht einwandfrei, denn es wird darin die Tatsache benötigt, daß die Originalreihe $P(t)$ schwach stationär sei. Andererseits unterstellen Granger und Hughes jedoch aufgrund der Daten ein Modell

$$P(t) = a(t) X(t),$$

in dem der Fluktuationsindex $X(t)$ schwach stationär mit $E(X(t)) = 1$ ist und $a(t)$ eine Trendfunktion, die durch $\bar{P}(t)$ geschätzt wird. Bei wachsenden $a(t)$ kann aber $P(t)$ nicht mehr schwach stationär sein.

19. Schätzprobleme beim Spektrum

19.1. Fragen zur Technik

Im Abschnitt 11 haben wir gesehen, daß die spektrale Dichtefunktion durch geglättete Periodogramme

$$(19.1) \quad \bar{f}_n(\lambda) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_n(\lambda - \beta) \bar{W}_m(\beta) d\beta = \frac{1}{2\pi} \bar{c}(0) + \frac{1}{\pi} \sum_{l=1}^m \bar{w}_m(l) \bar{c}(l) \cos \lambda l$$

konsistent geschätzt werden kann, wenn das Spektralfenster $\bar{W}_m(\lambda)$,

1) Vgl. König und Wolters (25, S.100)

dessen Fourierkoeffizienten $w_m(k)$ für $k > m$ gleich null sind mit n gegen unendlich gegen die Dirac-Funktion konvergiert. Gleichzeitig muß das Verhältnis zwischen der maximalen lag-Länge m und dem Stichprobenumfang n gegen Null gehen.

Bei der praktischen Anwendung muß entschieden werden, welches Spektralfenster zu wählen ist und wie groß bei fest vorgegebenen Stichprobenumfang die maximale lag-Länge m sein soll. Diese Fragen hat u.a. Naeve¹⁾ an einem autoregressiven Prozeß 2. Ordnung und an einem 6-gliedrigen gleitenden Durchschnitt untersucht. Dabei sind jeweils 300 Werte erzeugt und zur Schätzung der spektralen Dichte bei verschiedenen Alternativen benutzt worden. Naeve kommt im wesentlichen zu dem Ergebnis, daß die Wahl des Spektralfensters keinen großen Einfluß auf die Schätzergebnisse hat. Dagegen sind die Schätzungen sehr reagibel auf m . Aus den Schätzergebnissen kann man schließen, daß für den autoregressiven Prozeß ein Wert von m zwischen 20 und 30, d.h. ungefähr zwischen $n/7$ und $n/10$ optimal ist. Beim autoregressiven Prozeß wird die optimale lag-Länge mit $m=10$ sehr viel kleiner. Letzteres erklärt sich leicht durch die Tatsache, daß bei dem fünf-gliedrigen gleitenden Durchschnitt die Kovarianzfunktion $c(l)$ schon für $l \geq 6$ verschwindet. Eine generelle Regelung läßt sich hier also nicht finden. Die optimale Wahl von m hängt von dem zu schätzenden Spektrum ab. Bei der Schätzung der realen ökonomischen Reihe - es handelt sich um den monatlichen Bierausstoß in einzelnen Bundesländern der BRD - die eine Länge von $n = 222$ Werte haben, wird $m=40$ verwendet.

Naeve untersucht mittels Simulation auch noch, an wieviel Stellen das Spektrum zu schätzen ist, damit der Verlauf durch Hinzunahme neuer Punkte keine wesentliche Änderung erfährt. Den Ergebnissen nach sollte diese Anzahl mindestens gleich der maximalen lag-Zahl m sein. Theoretisch kann man auf das Doppelte schließen. Denn m ist die größte in der Fourierreihe (19.1) auftretende Frequenz; also ist $2\pi/m$ die

1) Naeve, P. (34, S.45 ff.)

kleinste Periode. In dem Bereich $0 \leq \lambda \leq \pi$ passen $m/2$ solcher Perioden. Setzt man zur ungefähren Beschreibung einer einzigen reinen Schwingung vier Punkte voraus, dann kommt man also auf $2m$ Stützstellen für die Schätzung des Spektrums.

19.2. Beurteilung einiger Schätzergebnisse

Wir wollen zunächst die Schätzergebnisse von Naeve etwas genauer betrachten.

Aus Beispiel 5.5 kennen wir für den gleitenden Durchschnitt folgende Beziehungen

$$X(t) = \sum_k \alpha_k U(t-k)$$

$$f(\lambda) = |a(\lambda)|^2 \frac{\sigma^2}{2\pi}, \quad a(\lambda) = \sum_k \alpha_k e^{-i\lambda k}$$

Dabei ist σ^2 die Varianz des white noise-Prozesses $U(t)$. Für den von Naeve simulierten Durchschnitt mit den Koeffizienten 0.1, 0.5, -0.1, 0.2, 0.1 errechnet man ($\sigma^2 = 1$)

$$\begin{aligned} |a(\lambda)|^2 &= \left(\sum_k \alpha_k e^{-i\lambda k} \right) \left(\sum_k \alpha_k e^{i\lambda k} \right) \\ &= \sum_k \alpha_k^2 + 2 \sum_{k > j} \alpha_k \alpha_j \cos \lambda(k-j) = f(\lambda)/2\pi \\ &= 0.36 + 0.25 \cos \lambda + 0.22 \cos \lambda 3 + 0.14 \cos \lambda 4 + \\ &\quad + 0.02 \cos \lambda 5 \end{aligned}$$

Daraus findet man für die spektrale Dichte und ihre zweite Ableitung die Abschätzungen

$$(19.1) \quad |f(\lambda)| \leq \frac{1}{2\pi}, \quad |f''(\lambda)| \leq \frac{5.11}{2\pi} \approx 2.$$

Bei dem von Naeve benutzten lag-Fenster von Tukey

$$w_m(l) = 0.5 + 0.5 \cos l \frac{\pi}{m}, \quad 0 \leq l \leq m,$$

bekommt man für den asymptotischen Schätzfehler

$$\begin{aligned} E(\bar{f}_n(\lambda) - f(\lambda))^2 &= B_n^2(\lambda) + \text{Var}(\bar{f}(\lambda)) \\ &\approx 0.0625 \left(\frac{\pi}{m}\right)^2 |f''(\lambda)|^2 + 0.75 \frac{m}{n} \end{aligned}$$

Mit (19.1) und den Werten $n=300$, $m=10$ ergibt sich eine obere Schranke

$$\begin{aligned} E(\bar{f}_n(\lambda) - f(\lambda))^2 &= 0.25 \frac{\pi^2}{100} + 0.75 \frac{1}{30} \\ &= 0.025 + 0.025 = 0.05. \end{aligned}$$

Die Schätzungen von Naeve¹⁾ liegen weit innerhalb des durch diese Schranke bestimmten Fehlerbereichs. Man kann also annehmen, daß bei diesem Beispiel - gleitender Durchschnitt $m=300$, $m=10$ - die asymptotischen Eigenschaften zur Beurteilung der Schätzung erfüllt sind. Die gleiche Aussage gilt für den untersuchten autoregressiven Prozeß. Das war auch theoretisch zu erwarten, denn die im Abschnitt 9 angegebenen asymptotischen Verzerrungen und Varianzen erhalten ihre Gültigkeit, wenn die Werte der Kovarianzfunktion $c(l)$ für $l > m$ vernachlässigt werden können und damit der fehlende Teil in der Schätzung des Spektrums auch tatsächlich keinen Einfluß mehr hat.

1) Vgl. (34, S.49, Graphik 3)

Daß dies für reale ökonomische Zeitreihen, die Trend- und Saisonkomponente enthalten, zutrifft, muß angesichts der Stationaritätsbetrachtungen im vorigen Abschnitt stark bezweifelt werden. Wenn man schon annimmt, diese Komponenten besäßen eine spektrale Dichtefunktion, dann muß man davon ausgehen, daß dieselbe starke Ausschläge in den Saisonfrequenzen und in der Nähe von $\lambda = 0$ besitzt. Die Frequenzausschläge werden verursacht durch periodisch wiederkehrende Korrelation nahe eins bzw. durch ein Überwiegen positiver Korrelationen überhaupt. In der Regel dürfte eine ökonomische Zeitreihe nicht lang genug sein, um den maximalen lag m so groß wählen zu können, daß alle wesentlichen Kovarianzen $c(1)$ erfaßt werden können. Selbst bei ausgesucht langen Reihen, wie etwa der schon erwähnte Bierausstoß, kann man das feststellen.¹⁾

Teilweise lassen sich Schätzfehler auch systematisch verfolgen. Ein bekannter Effekt ist das "Durchsickern" bei den Spektralfenstern. Dabei können Spitzen im Periodogramm durch die Seitenfenster im Schätzpunkt λ durchschlagen. Zur Illustration des entstehenden Schätzfehlers betrachten wir das Spektralfenster

$$W_m(\phi) = \frac{1}{2\pi} \frac{\sin^2 m\phi/2}{\sin^2 \phi/2}$$

Die Nullstellen der Funktion liegen bei $\phi_k = \pi k/m$, $k=2,4,\dots$. Die Maxima liegen zwischen diesen Nullstellen und können durch

$$\phi_k = \pi k/m, \quad k=0,3,5,\dots$$

approximiert werden.

1) Naeve (34, S.99 f.) Man beachte, daß die dort vorliegende fallende Tendenz der Korrelationsfunktion auf den vorhandenen linearen Trend zurückzuführen ist.

Im Nullpunkt erhält man durch Grenzübergang $\phi \rightarrow 0$:

$$W_m(0) = m^2/2\pi .$$

Für das erste Nebenmaxima erhalten wir approximativ

$$W_m(3\pi/m) = \frac{1}{2\pi} \frac{\sin^2 3\pi/m}{3\pi/2m}$$

$$m/3\pi^2 = 2W(0)/3\pi m .$$

Bei $m=10$ bekommen wir also ein Durchsickern von ca. 2% aus den Werten im ersten Seitenfenster. Der Effekt ist bei glattem Verlauf des Spektrums zwar gering, kann aber bei großen Spitzen, wie sie besonders vom Trend und von der Saisonkomponente verursacht werden, schon bedeutungsvoll sein.

Die Schätzfehler können Anlaß für Fehlinterpretationen sein. Setzt man z.B. voraus, daß die im Abschnitt 18 geschilderten Ergebnisse der Untersuchung von Granger und Hughes signifikant sind, das tatsächliche Spektrum der Weizenpreise also einen Gipfel für die Periode 13,3 und nicht für die Periode 15,0 aufzeigt, dann müßte aus den Graphiken bei Naeve (34, S.49 f.) auf bedeutende Gipfel an verschiedenen Stellen der spektralen Dichte geschlossen werden. In Wirklichkeit jedoch haben die Dichtefunktionen des autoregressiven Prozesses und des gleitenden Durchschnitts einen glatten Verlauf. Unter Berücksichtigung der asymptotischen Fehler kann man solche Fehlinterpretationen an einzelnen Stellen vermeiden. Dennoch fällt es schwer, den Gesamtverlauf der Schätzung mit dem theoretischen Verlauf in Einklang zu bringen.

Noch schwieriger ist die Beurteilung von geschätzten Spektren ökonomischer Reihen kürzerer Länge, wie sie in der Regel für wirtschaftliche Untersuchungen vorliegen. Zu diesem Zweck sind Simulationsstudien für

kurze Zeitreihen wiederum von Granger und Hughes (15) und auch von König und Wolters (24) und (25) angestellt worden. Diese zeigen zwar, daß für einfachste Prozeßtypen - white noise, autoregressiver Prozeß 1. Ordnung - kaum Verzerrungen und nur geringe Varianzen zu erwarten sind. Die Übertragbarkeit dieser Ergebnisse auf ökonomische Reihen bleibt jedoch zweifelhaft, denn bei Prozessen mit stärker ausgeprägtem Spektrum - fünfgliedriger gleitender Durchschnitt - waren die Ergebnisse nicht mehr zufrieden stellend.¹⁾

Vgl. König und Wolters (24, S. 159)

VI. Zusammenfassende Beurteilung und Konsequenzen

20. Kritik des Analysemodells

Im Abschnitt 4.1 haben wir als erstes Kriterium zur Beurteilung die Überprüfung genannt, ob die im Rahmen der ökonomischen Theorie zu erwartenden Komponentenformen in dem Modell integriert werden können. Für das Modell des schwach stationären Prozesses, der ergodisch ist und eine absolut stetige Spektralverteilung besitzt, muß dies verneint werden, denn typische Eigenschaften des Trends sprechen gegen schwache Stationarität und typische Eigenschaften der Saisonkomponente sprechen für Sprünge im Spektrum und damit gegen die Annahme einer spektralen Dichtefunktion.

Ein weiteres Argument gegen die Integration dieser beiden Komponenten ergibt sich aus einer anderen Überlegung. Dazu nehmen wir der Einfachheit halber an, der Prozeß X möge konstante Erwartungswerte gleich null besitzen. Dann sind auch die Erwartungswerte des Spektralprozesses Z konstant gleich null. Haben nun die Prozeßvariablen $Z(\lambda)$ eine zum Nullpunkt symmetrische Verteilung - bei einem Gaußschen Prozeß ist dies automatisch der Fall - dann gilt für eine Realisation $x = (\dots, x(-1), x(0), x(1), \dots)$

$$P(X=x) = P(X=-x)$$

d.h., die Wahrscheinlichkeit für den Verlauf einer Zeitreihe ist genau so groß, wie die Wahrscheinlichkeit für den Verlauf der negativen Reihe. Diese Erscheinung ist akzeptabel für die zyklische und die irreguläre Komponente. Bei jeder ökonomischen Zeitreihe kann man sich beide Komponenten in einem bestimmten Zeitabschnitt mit beliebigem Vorzeichen

vorstellen, sofern der Verlauf vor diesem Zeitabschnitt nicht bekannt ist. Für die Trend- und Saisonkomponente trifft dies aber nicht zu. Auch bei Nichtkenntnis der Vorgeschichte hat man bei diesen Komponenten für eine bestimmte Zeitreihe Anhaltspunkte über das Vorzeichen in einem gegebenen Zeitabschnitt. Insofern ist das stochastische Modell der Spektralanalyse zu unspezifiziert und allgemein. Die vorhandene Information über die Trend- und Saisonkomponente wird im Modell nicht aufgenommen.

Bezüglich der Sparsamkeit des Modells müssen ebenfalls Einwände gemacht werden. Die Menge der unbekannten beschreibenden Größen des Modells besteht aus der Menge der Zufallsvariablen $Z(\lambda)$, $-\pi = \lambda = \pi$, bzw. aus der Menge der Zuwächse $dZ(\lambda)$, die in der Spektraldarstellung

$$(20.1) \quad X(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ(\lambda)$$

vorkommen. Da die Zuwächse orthogonal sind, der einzelne Beitrag von $dZ(\lambda)$ in (20.1) aber verschwindend gering ist, muß die Beschreibung des schwach stationären Prozesses mittels der Spektraldarstellung als sehr aufwendig angesehen werden. Es kommt noch hinzu, daß den Zuwächsen keine ökonomische Bedeutung gegeben werden kann.

Die Einwände bezüglich der Sparsamkeit schließen natürlich die Annahme der schwachen Stationarität für die zyklische und die irreguläre Komponente nicht aus. Es scheint aber angebracht zu sein, diese in einer anderen Form darzustellen. Dazu bieten sich wegen des stetigen Spektrums die gleitenden Durchschnitte und autoregressiven Prozesse an.

Der Gedanke, andere Modelle in Betracht zu ziehen, wird auch nahegelegt, wenn wir uns die Theorie zur linearen Prognose aus Kapitel III in Erinnerung rufen. Zunächst stellt man fest, daß ökonomische Reihen sicherlich als Realisationen von regulären Prozessen anzusehen sind, d.h.: Ist $X(t,1)$ die Prognose vom Zeitpunkt $t-1$ für den Zeitpunkt t

unter Berücksichtigung der gesamten Vergangenheitswerte $X(t-1)$, $X(t-2)$, ... dann gilt

$$\sigma^2(1) = \text{Var} (X(t,1) - X(t)) > 0 .$$

Trifft dies aber zu, dann kann nach der Zerlegung von Wold in Abschnitt 12 der Prozeß X als Summe

$$X(t) = \sum_{k=0}^{\infty} \alpha_k U(t-k) + V(t) , \quad \sum \alpha_k^2 < \infty , \quad \alpha_0 = 1$$

geschrieben werden. Dabei entspricht der deterministische Prozeß $V(t)$ dem singulären Teil in der Spektralverteilung und der gleitende Durchschnitt dem absolut stetigen Teil.

Auch die zyklische Komponente und die Restkomponente können ohne Einwände als regulär betrachtet werden. Da sie ein stetiges Spektrum besitzen, sind sie identisch mit ihrer Wold-Zerlegung und also zusammen betrachtet ein einseitiger gleitender Durchschnitt. Hier nun können die Theorie ökonometrischer Modelle und die Spektraltheorie in eine fruchtbare Verbindung gebracht werden, indem wir den in der Gleichung (1.5) auftretenden stochastischen Teil

$$Y(t) = \sum_{\tau=0}^{t-1} A^{\tau} V(t-\tau)$$

des Verlaufs der endogenen Variablen eines ökonometrischen Modells zur Erklärung der zyklischen Schwankungen heranziehen. Für unbeschränkt¹⁾ wachsendes t handelt es sich um den gleitenden Durchschnitt eines Vektorprozesses. Jede Komponente ist aufgrund der Spektraltheorie zu zyklischen Schwankungen fähig, deren mehr oder

1) Da die Eigenwerte von A alle kleiner als 1 sein sollen, kann der Grenzübergang $t \rightarrow \infty$ vollzogen werden.

weniger starke Ausprägungen von der spektralen Dichtefunktion abhängen. Die wiederum läßt sich aus den Koeffizienten des gleitenden Durchschnitts errechnen. Über die Schätzung der Koeffizientenmatrizen des ökonometrischen Modells bekommt man also auch Aussagen über die durch den stochastischen Teil verursachten zyklischen Schwankungen. Untersuchungen dazu hat Howrey (20) am bekannten Klein-Goldberger-Modell angestellt. In einer unveröffentlichten Arbeit von Laven (30) ergeben sich aus der Schätzung eines einfachen Modells für die BRD in einigen Fällen Spektren, die auf Schwankungen mit einer Periode von 5-6 Jahren schließen lassen, wie sie in der Realität tatsächlich beobachtet wurden.

21. Kritik der Ziele

Das Hauptanliegen der Spektralanalyse war die Bestimmung der spektralen Dichtefunktion. Gemessen am Bestreben der Zeitreihenanalyse schlechthin, wonach die numerischen Werte der Komponenten bestimmt werden sollen, bleibt jenes Anliegen weit zurück; denn aus der Kenntnis der spektralen Dichtefunktion kann man nicht auf den Verlauf einer Realisation des schwach stationären Prozesses schließen. Die Dichte bzw. die Verteilungsfunktion enthält nämlich nur diejenige Information, die auch die Kovarianzfunktion enthält. Das ist eine Konsequenz aus den Gleichungen

$$f(\lambda) = \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} c(l) e^{-i\lambda l}, \quad c(l) = \int_{-\pi}^{\pi} e^{i\lambda l} f(\lambda) d\lambda,$$

wonach die $c(l)$ die Fourierkoeffizienten der spektralen Dichte sind. Allerdings ist die Wirkung - oder besser gesagt eine bestimmte Wirkung - einer gewissen Konstellation von Kovarianzen in der Dichte-

funktion bzw. in der spektralen Verteilungsfunktion besser zu erkennen. Das ist möglich wegen des Zusammenhangs

$$E(|dZ(\lambda)|^2) = dF(\lambda)$$

der zwischen Spektralprozeß $Z(\lambda)$ und Verteilungsfunktion $F(\lambda)$ besteht. Aus dieser Beziehung kann man ablesen: Konzentration spektraler Masse in einem bestimmten Frequenzbereich bedeutet, daß in diesem Bereich die Variablen des Spektralprozesses eine große Varianz aufweisen. Je größer diese Varianzen sind, desto wahrscheinlicher werden die Realisationen größer als sonst ausfallen. Dies bewirkt, daß bei einer Realisation des stochastischen Prozesses X mit entsprechenden periodischen Bewegungen zu rechnen ist, falls entsprechend große $dZ(\lambda)$ realisiert wurden. Die Überlegungen dürfen aber nicht zu der Annahme verleiten, daß die Werte des Periodogramms

$$P_n(\lambda) = \frac{1}{4\pi n} \left| \sum_{t=1}^n e^{i\lambda t} X(t) \right|^2$$

an den Stellen $\lambda_k = 2\pi k/n$, $0 \leq k \leq n/2$ proportional zu den Amplitudenquadraten eine Fourierzerlegung mit diesen Frequenzen sind. Wir sehen also, daß eine Berechnung bestimmter Zeitreihenkomponenten - die sich aus den Amplituden der Fourierzerlegung sofort ergibt - über das Periodogramm nicht möglich ist.

Dem Ziel der numerischen Bestimmung der Komponenten kommen die Filtermethoden wesentlich näher. Allerdings muß man hier genau unterscheiden zwischen der Wirkung der Filterfunktion $b(\lambda)$ und der Wirkung der Übertragungsfunktion $|b(\lambda)|^2$. Es sei z.B. für einen Frequenzbereich I approximativ die Bedingung

$$|b(\lambda)|^2 = \begin{cases} 0 & \text{für } \lambda \in I \\ 1 & \text{sonst} \end{cases}$$

erfüllt. Dann hat zwar der gefilterte Prozeß

$$Y(t) = \sum_k \beta_k X(t-k)$$

die gewünschte spektrale Dichtefunktion, die in I verschwindet und außerhalb mit der spektralen Dichte von X übereinstimmt. Wegen der Beziehung

$$dZ_y(\lambda) = b(\lambda) dZ_x(\lambda)$$

gilt die entsprechende Aussage für die Spektralschar jedoch erst, wenn auch approximativ

$$(21.1) \quad b(\lambda) = \begin{cases} 0 & \text{für } \lambda \in I \\ 1 & \text{sonst} \end{cases}$$

erfüllt ist. Bei reellwertigen Prozessen müssen deswegen die reellen Filterkoeffizienten die Symmetrieeigenschaft $\beta_k = \beta_{-k}$ erfüllen. Wenn wir also als Ziel der Komponentenausschaltung formulieren, daß der schwach stationäre Prozeß in zwei Anteile zerlegt werden soll, die in zueinander komplementären Bereichen die Spektralschar von X besitzen, dann ist dies im Grunde nur mit zu niedrigen Filtern möglich. Bei genügend guter Approximation von (21.1) - die allerdings durch die Länge der Zeitreihe sehr beschränkt ist - kann $Y(t)$ als der bereinigte Prozeß und $X(t)-Y(t)$ als die ausgeschaltete Komponente angesehen werden.

In diesem Zusammenhang wird ein Mangel deutlich, der die Definition der einzelnen Komponenten betrifft, wie sie im stationären Modell integriert wurden. Nehmen wir an, eine Zeitreihe sei nach den obigen Überlegungen von Trend-, Saison- und zyklischen Einflüssen bereinigt worden, dann wäre der Rest - die irreguläre Komponente - ein Prozeß, dessen spektrale Dichte in gewissen Frequenzbereichen gleich null ist. Andererseits müßte aber diese Komponente eine konstante Dichtefunktion besitzen.

22. Kritik der Schätzmethoden

Wir unterscheiden auch hier wieder zwischen der Filtermethode zum Zwecke der Ausschaltung von Komponenten und der Schätzung der spektralen Dichtefunktion. Zunächst müssen wir feststellen, daß die Rechnungen der Filtermethode genauso zu beurteilen sind, wie diejenigen der Methode des gleitenden Durchschnitts in Abschnitt 4.2. Der Unterschied zwischen beiden Methoden besteht darin, daß wir für das Filterverfahren ein differenzierteres Analysemodell zur Verfügung haben. Aufgrund dieses Modells sind wir in der Lage den Grund dafür anzugeben, weswegen die Rechnungen unbefriedigend sind. Er besteht vor allem in der Kürze ökonomischer Zeitreihen, die verhindern, daß eine Filterfunktion, die der asymptotischen Eigenschaft (21.1) nahekommt, angewendet werden kann. Wie uns das Kapitel III gezeigt hat, sind Filtermethoden auch zur Prognose geeignet.

Die Schätzung der spektralen Dichtefunktion kann natürlich nicht an den Kriterien von Lovell¹⁾ gemessen werden, weil keine Bestimmung von Komponenten stattfindet. Die Schätzung selbst ist, wie wir in Kapitel IV gesehen haben, konsistent. Sie besitzt allerdings schlechte asymptotische Eigenschaften. Bei allen in Abschnitt 15.2 betrachteten Spektralfenstern sind die Varianzen und die Verzerrungen von der Größenordnung

$$\text{Var}(\bar{f}_n(\lambda)) = O(m/n) \quad , \quad B_n(\lambda) = O(1/m)$$

wobei diese Beziehungen gelten, wenn die maximale Zahl m der geschätzten Kovarianzen mit dem Stichprobenumfang n gegen unendlich geht, gleichzeitig aber m/n gegen 0 konvergiert. Die Konvergenz

1) Vgl. Abschnitt 4.2

des Schätzfehlers

$$E(\bar{f}_n(\lambda) - f(\lambda))^2 = B_n^2(\lambda) + \text{Var}(\bar{f}(\lambda))$$

ist deshalb auf jeden Fall schlechter als $1/n$. Eine genauere Betrachtung findet man bei Grenander und Rosenblatt (17, S. 153ff.).

Diese Tatsache, gepaart mit dem geringen Stichprobenumfang bei ökonomischen Reihen, scheint die Schätzung spektraler Dichten zur Feststellung signifikanter Konzentrationen wertlos zu machen. Aus dem Studium vieler Schätzungen glaubt aber Granger (13) zumindest das Ergebnis ablesen zu können, daß die spektralen Dichten vieler ökonomischer Zeitreihen nach einer Trend- und Saisonbereinigung eine typische Figur aufweisen, die auf einen einfachen autoregressiven Prozeß

$$X(t) = \beta X(t-1) + U(t), \quad 0 < \beta < 1,$$

zurückgehen. Die Dichtefunktionen von X und dem white noise Prozeß U stehen nämlich in der Beziehung¹⁾

$$|1 - \beta e^{-i\lambda}| f_X(\lambda) = f_U(\lambda) = \sigma_u^2/2\pi,$$

woraus sich

$$f_X(\lambda) = \sigma_u^2(1 + \beta^2 - 2\beta \cos \lambda)/2\pi$$

ergibt. Diese Funktion hat für β in der Nähe von eins einen hyperbelartigen Verlauf, wie ihn Granger im Sinn hat. Wenn auch das von Granger zur Illustration benutzte Beispiel eines Multiplikator-Akzelerator-Modells sehr einfach ist, so bestätigt er doch mit seinen Überlegungen, daß die Schätzung der spektralen Dichte von ökonomischen Zeitreihen nicht als Selbstzweck betrieben werden kann, sondern, daß erst die Hinzunahme anderer Formen stationärer Prozesse

1) Vgl. Beispiel 9.7

ökonomische Interpretationen der Ergebnisse ermöglicht.

23. Alternativen zur Spektralanalyse

Bei grober Vereinfachung können wir der Untersuchung entnehmen:

- a) Mit Trend- und Saisoneinflüssen behaftete ökonomische Reihen sind im allgemeinen weder stationär, noch besitzen sie ein stetiges Spektrum. Dagegen sprechen starke Argumente dafür, diese Eigenschaften für die zyklische und die irreguläre Komponente vorauszusetzen.
- b) Spektralanalyse als statistische Methode zur Schätzung des Spektrums ist, von Ausnahmen abgesehen, bei ökonomischen Reihen wegen der fehlenden Signifikanz i.a. wertlos.

Als Möglichkeiten der univariaten Zeitreihenanalyse, die dieser Situation Rechnung tragen, möchten wir noch zwei Alternativen aufzeigen.

23.1. Ein Regressionsmodell mit stationärem "Rest"

Wir haben im Abschnitt 3 ein lineares Regressionsmodell betrachtet. Indem wir die dort benutzte Matrix der Ansatzfunktionen A reduzieren auf die Spaltenvektoren, die zur Beschreibung von Trend- und Saisonkomponenten benötigt werden und die zyklische Komponente mit der irregulären zusammenfassen, erhalten wir ein Regressionsmodell mit einer stochastischen Komponente V , die bei Berücksichtigung der obigen Argumente stationär ist und die

Varianz-Kovarianz-Matrix

$$\Gamma = E(VV') = \begin{bmatrix} c(0) & c(1) & \dots & c(n) \\ c(1) & c(0) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ c(n) & \dots & \dots & c(0) \end{bmatrix}$$

besitzt. Bekanntermaßen ist die beste lineare unverzerrte Schätzung des Regressionsmodells in diesem Fall gegeben durch

$$\hat{\beta} = (A' \Gamma^{-1} A)^{-1} A' \Gamma^{-1} X$$

Der Wert kann aber nicht berechnet werden, da die Matrix Γ unbekannt ist. Man kann aber die Schätzung nach der Methode der kleinsten Quadrate

$$\hat{\beta} = (A' A)^{-1} A' X$$

bestimmen und nach der Effizienz fragen. Es läßt sich zeigen¹⁾, daß $\hat{\beta}$ asymptotisch die gleiche Varianz besitzt wie $\hat{\beta}$, wenn für den Restprozeß V eine spektrale Dichtefunktion existiert, die stückweise stetig ist und die Ansatzfunktion der Matrix A gewisse Bedingungen erfüllen, die bei Polynomen und trigonometrischen Funktionen, wie sie für Trend- und Saisonkomponente benötigt werden, erfüllt sind. Da auch die Annahme über die spektrale Dichtefunktion von V gerechtfertigt ist, kann man diesen Regressionsansatz als geeignet dafür ansehen, die systematischen Komponenten $E(X) = A\beta$ einer ökonomischen Zeitreihe zu schätzen, wobei gleichzeitig die stochastischen Eigenschaften der zyklischen Komponente voll berücksichtigt werden.

1) Vgl. Grenander und Rosenblatt (17, S.233 ff.) oder Anderson, T.W. (3, S. 900 ff.).

23.2 ARMA und ARIMA-Modelle

Wir betrachten noch eine geeignete Alternative für die Untersuchung einer ökonomischen Reihe $V(t)$, die stationär ist. Das betreffende Modell lautet:

$$\begin{aligned} V(t) + b_1 V(t-1) + \dots + b_p V(t-p) &= \\ &= U(t) + a_1 U(t-1) + \dots + a_q U(t-q) . \end{aligned}$$

Dabei ist $U(t)$ ein weißes Rauschen. Box und Jenkins (6) bezeichnen dies als ein "ARMA" (autoreressive moving average)-Modell. Sind die Wurzeln des Polynoms

$$z^p + b_1 z^{p-1} + \dots + b_p = 0$$

dem Betrage nach alle kleiner als eins, dann ist der Prozeß stationär und kann als gleitender Durchschnitt

$$V(t) = \sum_{k=0}^{\infty} \alpha_k U(t-k) , \quad \alpha_0 = 1 ,$$

dargestellt werden¹⁾. Erfüllen die Wurzeln des Polynoms

$$z^q + a_1 z^{q-1} + \dots + a_q = 0$$

die gleiche Bedingung, dann ist $V(t)$ auch ein autoregressiver Prozeß

$$V(t) = \sum_{k=1}^{\infty} \beta_k V(t-k) + U(t) .$$

Mit diesen beiden Voraussetzungen über die charakteristischen Polynome ist das ARMA-Modell für die Analyse der zyklischen Komponente und der Restkomponente geeignet. Es enthält alle die Eigenschaften, die wir über diese Komponenten festgestellt haben und bietet zusätzlich den

1) Vgl. die Beispiele 9.6 und 9.7.

Vorteil, nur durch wenige Parameter erklärt zu sein. Box und Jenkins kommen bei ihren Schätzungen i.a. mit $p, q = 2$ aus. Hinzu kommt noch die zu bestimmende Varianz von $U(t)$. Das Modell erlaubt außerdem, wegen der autoregressiven Darstellung auf einfachem Wege Prognosen zu bestimmen.

Für den Fall, daß eine Zeitreihe $X(t)$ nicht stationär ist, führen Box und Jenkins eine Methode ähnlich der Variate Difference-Methode ein. Es werden von der Ausgangsreihe solange Differenzen gebildet, bis aufgrund bestimmter Kriterien über die Kovarianzen - die Kriterien sind andere als diejenigen der Variate Difference-Methode - Stationarität angenommen werden kann. Das Ergebnis ist das $V(t)$ des ARMA-Modells. Ist ∇ die Differenzenbildung

$$\nabla X(t) = X(t) - X(t-1)$$

dann erhält man also nach d Schritten

$$V(t) = \nabla^d X(t) .$$

Unter dieser Berücksichtigung wird aus dem ARMA-Modell ein sogenanntes ARIMA-Modell ($I = \text{integrated}$) mit dem zusätzlichen Parameter d , der ebenfalls auf kleine Werte beschränkt werden kann. Statt der Differenzen aufeinanderfolgender Werte können auch z.B. Differenzen

$$\nabla_{12} X(t) = X(t) - X(t-12)$$

gebildet werden, wenn Saisoneinflüsse auszuschalten sind. Ohne auf die Problematik der Methoden einzugehen, möchten wir festhalten, daß die ARMA-Modelle in Verbindung mit Regressionsmodellen im Rahmen der univariaten Analyse ökonomischer Reihen sicherlich zu den Analysemodellen zählen, welche die besten Ergebnisse liefern.

Literaturverzeichnis

- (1) Adelmann, I. und Adelmann, F.L.
The Dynamic Properties of the Klein-Goldberger
Model, *Econometrica* 28 (1959), S. 596-625
- (2) Anderson, O.
Zur Problematik der empirisch-statistischen
Konjunkturforschung, Veröffentlichungen der
Frankfurter Gesellschaft für Konjunkturforschung
Heft 1, Bonn 1929
- (3) Anderson, T.W.
The statistical Analysis of Time Series,
John Wiley and Sons 1971
- (4) Bartlett, M.S.
Periodogram Analysis and Continuous Spectra,
Biometrika 37 (1950), S. 1-16
- (5) Blackmann, R.B. und J.W. Tuckey
The Measurement of Power Spectra,
New York 1958
- (6) Box, G.E.P. und G.M. Jenkins
Time Series Analysis Forecasting and Control
Holden Day San Francisco 1970
- (7) Doob, J.L.
Stochastic Prozesses, John Wiley and Sons
New York 1967
- (8) Durbin, J. und G.S. Watson
Testing for Serial correlation in Least Squares
Regression I, *Biometrika* 37 (1950), S. 409-428
- (9) Durbin, J. und G.S. Watson
Testing for Serial correlation in Least Squares
Regression II, *Biometrika* 38 (1951), S. 159-178
- (10) Fries, C.
Die Ausprägungen der Spektraltheorie in Zeitreihen-
analyse und Quantenmechanik, *Statistische Hefte* 1969,
S. 271
- (11) Gandolfo, G.
Mathematical Methods and Models in Economic Dynamics,
Amsterdam 1962

- (12) Gerthsen, Ch.
Physik, Springer Berlin 1963
- (13) Granger, C.W.J.
The typical Spectral Shape of an Econometric Variable, *Econometrica* 34 (1966) S. 150-161
- (14) Granger, C.W. J. und M. Hatanaka
Spectral Analysis of Economic Time Series,
Princeton, New Jersey 1964
- (15) Granger, C.W.J. und A.O. Hughes
Spectral Analysis of Short Series - A Simulation Study, *Journal of Roy.Stat. Soc., Series A* 131 (1968) S. 83-89
- (16) Granger, C.W.J. und A.O. Hughes
A new Look at Old Data: Beveridge Wheat Price Series
Journal of Roy.Stat.Soc., Series A 131 (1971)
- (17) Grenander, U. und M. Rosenblatt
Statistical Analysis of Stationary Time Series,
New York 1957
- (18) Hannan, E.J.
Time Series Analysis, London 1960
- (19) Hannan, E.J.
The Estimation of Seasonal Variation in
Economic Time Series, *Journal of American Stat. Assoc.*, Vol. 58 (1963) S. 31-44
- (20) Howrey, E.P.
Stochastic Properties of the Klein-Goldberger Model
Econometrica 29 (1971)
- (21) Jenkins, G.M. und D.G. Watts
Spectral Analysis and its Applications, London 1968
- (22) Kendall, M.G. und Stuart
The Advanced Theory of Statistics,
Volume 3, London 1966
- (23) Kintchine, A.
Korrelationstheorie der stationären stochastischen
Prozesse, *Math. Ann.*, Band 109 (1934)

- (24) König, H. und J. Wolters
Spektralschätzungen stationärer stochastischer
Prozesse: Eine Simulationsstudie, Jahrbuch
für Nationalökonomie und Statistik, Band 185 (1971)
S. 142-162
- (25) König, H. und J. Wolters
Einführung in die Spektralanalyse ökonomischer
Zeitreihen, Meisenheim a. Glan 1972
- (26) Kolmogorov, A.N.
Sur l'interpolation et l'extrapolation des suites
stationnaires, Compt. rend. acad. sci., Paris,
Vol. 208 (1939) S. 2043-2045
- (27) Kondratieff, N.D.
Die langen Wellen der Konjunktur, Archiv für Sozial-
wissenschaft und Sozialpolitik, Band 56 (1926) S.574
- (28) Krelle, W.
Beeinflußbarkeit und Grenzen des Wachstums,
in: Wachstum und Entwicklung der Wirtschaft,
Herausgeber H. König, Köln 1968
- (29) Krickeberg, K. Wahrscheinlichkeitstheorie,
Teubner Stuttgart 1963
- (30) Laven, G.
Unveröffentlichtes Manuskript
- (31) Lovell, M.C.
Seasonal Adjustment of Economic Time Series and
Multiple Regression Analysis, Journal of Am.
Stat. Assoc., Volume 58 (1963) S. 994
- (32) Matthews, R.C. O.
Konjunktur, Nymphenburger Verlag, München (1973)
- (33) Morgenstern, O.
A new Look of Economic Time Series Analysis
in: Money, Growth, and Methodology and other
Essays in Economics, Edited by Hegeland, H.,
Lund 1961
- (34) Naeve, P.
Spektralanalytische Methoden zur Analyse von
ökonomischen Reihen, Würzburg 1969

- (35) Nerlove, M.
Spectral Analysis of Seasonal Adjustment Procedures,
Econometrica Volume 32 (1964) S. 241-286
- (36) Nullau, B.
Verfahren zur Zeitreihenanalyse, Vierteljahreshefte
zur Wirtschaftsforschung 1968, Heft 1
- (37) Parzen, E.
On Asymptotically Efficient Consistent Estimates
of the Spektral Density Function of Stationary
Time Series, Journal of Roy.Stat.Soc., Series B
(1958), S. 303-322
- (38) Philipps, A.W.
A simple Model of Employment, Money and Prices
in a Growing Economy, Economica 28 (1961) S. 360-370
- (39) Riesz, F. und B. S.Z. Nagy
Vorlesungen über Funktionalanalysis, Berlin 1968
- (40) Schäffer, K.
Beurteilung einiger herkömmlicher Methoden zur
Analyse von ökonomischen Zeitreihen,
Sonderheft im allgem. Stat. Archiv, Göttingen 1970
- (41) Schumpeter, J.A.
Mitchels Business Cycles, Quarterly Journal of
Economic, 45 (1930-31) S. 150-172
- (42) Schumpeter, J.A.
Konjunkturzyklen Göttingen 1961
- (43) Schuster, Sir Arthur
On the Investigation of Hidden Periodicities. With
Application to a Supposed Twenty Six Day Period of
Meteorological Phenomena, Terrestrial Magnetism
Volume 3 (1898), S. 13
- (44) Schiskin, J.
The X-11 Variant of the Census Method II Seasonal
Adjustment Program, U.S. Bureau of the Census,
Technical Paper No. 15, Washington DC 1965
- (45) Slutsky, E.
The Sumation of Random Causes as the Source of
Cyclical Processes, Econometrica, Volume 5,
(1937) S. 105-146

- (46) Smirnov, W.J.
Lehrgang der höheren Mathematik Band V, Verlin 1967
- (47) Urbanik, K.
Lectures on prediction theory, Berlin 1967
- (48) Wald, A.
Berechnung und Ausschaltung von Saisonschwankungen,
Wien 1936
- (49) Weinstock, U.
Das Problem der Kondratieff-Zyklen,
Berlin-München 1964
- (50) Whittle, P.
Hypotheses Testing in Time Series Analysis,
Uppsala 1951
- (51) Wiener, N.
Generalized Harmonic Analysis, Acta Math.,
Volume 35 (1930), S. 117-258
- (52) Wiener, N.
Extrapolation, Interpolation and Smoothing of
Stationary Time Series, Technology Press,
Cambridge, and John Wiley and Sons, New York 1949
- (53) Wold, H.
A Study in the Analysis of Stationary Time Series,
Almquist and Wicksells, Uppsala 1954
- (54) Yule, G.U.
On a Method of Investigating Periodicities
in Disturbed Series with Special Reference to
Wolfers' Sunspot Numbers, Phil. Trans. Roy.Soc.,
Volume 226, Series A (1927) S. 267-298

L e b e n s l a u f

Am 22. Dezember 1941 wurde ich als fünftes Kind des Landwirts Ferdinand Oswald und seiner Ehefrau Hilda geb. Happ in Worms geboren.

Ich besuchte ab Herbst 1948 die Volksschule in Gimbsheim Kreis Worms und ab Ostern 1953 das Gauß-Gymnasium in Worms, wo ich 1962 die Reifeprüfung ablegte.

Im Wintersemester 1962/63 studierte ich Physik an der Universität Mainz. Nach einer einjährigen Dienstzeit bei der Bundeswehr begann ich im Sommersemester 1964 das Mathematikstudium in Mainz, das ich 1970 mit der Diplomprüfung abschloß.

Seit April 1970 bin ich wissenschaftlicher Mitarbeiter im Fachbereich Rechts- und Wirtschaftswissenschaften der Johannes Gutenberg-Universität. Gleichzeitig war ich bis zum Sommersemester 1973 ordentlicher Student der Fachrichtung Volkswirtschaftslehre.

Die mündliche Doktorprüfung bestand ich am 20. Dezember 1976.

Mainz, den 24. Januar 1977

Eugen Oswald

Dissertationsdruck
GmbH & Co. KG.

M. SCHADEL

86 Bamberg · Hohe-Kreuz-Str. 9b · Tel. 0951/21617[5 4414]